DOCUMENT RESUME

ED 408 846 FL 024 604

| | |
|---|---|
| AUTHOR | Auger, Julie; Rose, Yvan |
| TITLE | Explorations du lexique (Explorations in Lexicon). Publication B-208. |
| INSTITUTION | International Center for Research on Language Planning, Quebec (Quebec). |
| REPORT NO | ISBN-2-89219-261-7 |
| PUB DATE | 97 |
| NOTE | 322p. |
| PUB TYPE | Books (010) -- Collected Works - General (020) |
| LANGUAGE | French |
| EDRS PRICE | MF01/PC13 Plus Postage. |
| DESCRIPTORS | American Indians; *Bilingualism; Child Language; Code Switching (Language); Contrastive Linguistics; Dictionaries; English; Food; Foreign Countries; French; Interlanguage; Language Research; Language Role; Language Usage; *Languages for Special Purposes; Linguistic Borrowing; Linguistic Theory; Morphology (Languages); Phonology; *Pragmatics; Pronunciation; Regional Dialects; *Semantics; Sociocultural Patterns; Sociolinguistics; Translation; Verbs; *Vocabulary; *Vocabulary Development |
| IDENTIFIERS | Armenia; Canada; French (Cajun); *French (Canadian); Kinyarwanda |

ABSTRACT

Essays on lexicon and vocabulary development, all in French, address these topics: bilingual lexicon in the context of translation; lexical and conceptual representation in bilinguals; pronouns in "existential" constructions in French and English; configurational functions within grammar; Gustaf Stern's diachronic semantic schema; the semantic categories of "event" and "action"; pragmatic connectors as indices of linguistic development of 5- to 11-year-olds; vocabulary needed by entering university students; disparate lexicalization in a minority context; semantic patterns in Montreal (Canada) anglophones and francophones; "francismes" in Canadian usage; English borrowing in Cajun French; computerized context analysis of English loan words in Canadian French; bilingual code-switching in an Armenian/French contact situation; Amerindianisms in Canadian French; phonemic preservation in the adaptation of loan words; English phonemes in Quebec (Canada) French; pronunciation of English loan words in Quebec French, particularly among anglophones; phonology and morphology in French loans to Kinyarwanda; phenomena related to new technical terms before they appear in the dictionary; criteria for selection of terms for a bilingual dictionary; inventory of food-related vocabulary; adaptation of French children's dictionaries for use in Quebec; and characteristics of sociopolitical vocabulary in Quebec. Each chapter contains references. (MSE)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED 408 846

dral / clip
CENTRE INTERNATIONAL DE RECHERCHE EN AMÉNAGEMENT LINGUISTIQUE
INTERNATIONAL CENTER FOR RESEARCH ON LANGUAGE PLANNING

# Explorations du lexique

Sous la direction de

***Julie Auger et Yvan Rose***

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Claude Rochelcau
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

**Publication B-208**

BEST COPY AVAILABLE

FACULTÉ DES LETTRES
UNIVERSITÉ LAVAL

2

**1997**

# Explorations du lexique

Sous la direction de

***Julie Auger et Yvan Rose***

**B-208**

1997
CENTRE INTERNATIONAL DE RECHERCHE EN AMÉNAGEMENT LINGUISTIQUE
INTERNATIONAL CENTER FOR RESEARCH ON LANGUAGE PLANNING
QUÉBEC

**Données de catalogage avant publication (Canada)**

Vedette principale au titre :

Explorations du lexique

(Publication B ; 208)
Textes d'un colloque présenté lors du Congrès de l'Association canadienne-française pour l'avancement des sciences, tenu à l'Université McGill en 1996.
Comprend des réf. bibliogr.

ISBN 2-89219-261-7

1. Lexicographie - Congrès. 2. Lexicologie - Congrès. 3. Emprunts (Linguistique) - Congrès. 4. Mot (Linguistique) - Congrès. 5. Sociolinguistique - Congrès. 6. Sémantique - Congrès. I. Auger, Julie. II. Rose, Yvan, 1971- . III. Centre international de recherche en aménagement linguistique. IV. Association canadienne-française pour l'avancement des sciences. Congrès (64e : 1996 : Université McGill). V. Collection : Publication B (Centre international de recherche en aménagement linguistique) ; 208.

P327.E96 1997 413'.028 C97-940210-7

© **CENTRE INTERNATIONAL DE RECHERCHE EN AMÉNAGEMENT LINGUISTIQUE**
*Tous droits réservés. Imprimé au Canada.*
Dépôt légal (Québec) – 1er trimestre 1997
ISBN: 2-89219-261-7

# *Préface*

Pour la première fois de son histoire, l'*Association canadienne-française pour l'avancement des sciences* (*ACFAS*) a tenu son congrès annuel à l'Université McGill. C'est dans ce cadre que s'est déroulé un colloque intitulé *Explorations du lexique*. Ce colloque, organisé par Julie Auger et des collègues des universités McGill et de Sherbrooke, a réuni près de cinquante chercheurs du Canada et des États-Unis.

Suite à la tenue du colloque, une proposition de publication arbitrée a été soumise aux auteurs des communications présentées lors du colloque. Ce livre regroupe les textes qui ont été acceptés par le comité d'arbitrage et est divisé selon cinq sections thématiques: la représentation mentale du lexique bilingue, la sémantique lexicale et la pragmatique, les aspects sociolinguistiques du lexique, les emprunts et les lexiques spécialisés.

La section sur la représentation mentale du lexique bilingue se compose de deux articles. Le premier s'interroge sur la pertinence d'utiliser la traduction comme méthode expérimentale. Cet article démontre la présence d'un lien direct entre le niveau de bilinguisme et la représentation lexicale. Le second article discute d'un modèle de représentation du lexique bilingue. On y soutient l'hypothèse que les bilingues ont à leur disposition un système de représentations mentales non linguistiques et deux systèmes lexicaux.

Dans la section sur la sémantique lexicale et la pragmatique, les cinq articles soulignent l'importance du mot. D'une part, il est démontré qu'en dépit de la diversité qui caractérise les mots dans leurs emplois et dans leur évolution, il demeure possible de proposer un sens de base que l'on peut altérer selon des principes systématiques (par exemple, *y* dans *Il y a un Père Noël*; *tiens!* qui garde un lien avec le verbe *tenir*). D'autre part, on s'interroge sur la forme que doivent prendre les entrées lexicales: on a évalué la nécessité de distinguer les catégories "événement" et "action" dans l'analyse des

verbes et démontre que, dans une théorie générative de la grammaire, une place importante doit être faite à des entrées lexicales bien structurées. Finalement, un article porte sur le développement de la cohérence textuelle chez les enfants à travers l'emploi de connecteurs.

La section sur les aspects sociolinguistiques comporte quatre articles. On s'y interroge d'abord sur la nature du vocabulaire testé dans les tests d'admission à l'université. Deux autres études ont pour objet des situations de contact linguistique entre le français et l'anglais. Ces études, qui portent sur le français en Ontario et sur l'apprentissage du français par des anglophones de Montréal illustrent de belle façon les avantages des études sociolinguistiques détaillées pour l'étude du lexique. Finalement, la notion de francisme, telle qu'utilisée dans les dictionnaires québécois récents, est remise en question.

La section sur les emprunts comprend huit articles portant sur une variété de situations de contact de langues. Si plusieurs travaux portent sur les adaptations orthographiques, phonologiques et grammaticales subies par les emprunts dans diverses langues, un texte rapporte l'absence d'adaptation des emprunts anglais en français cadjin. Un autre article se concentre sur les problèmes que posent les mots anglais en français aux anglophones qui apprennent le français. Finalement, un article porte sur la désambiguïsation automatique des emprunts dans des banques de données textuelles.

La section finale de ce collectif se penche sur les lexiques spécialisés. Les cinq articles de cette section portent sur des questions d'ordre lexicographique. La problématique centrale concerne la rédaction de dictionnaires québécois et canadiens, unilingues et bilingues: sélection des termes, rédaction de définitions, ainsi que le choix des exemples. Les problèmes de nature linguistique et culturelle que pose l'adaptation d'un dictionnaire d'apprentissage du français y sont abordés.

Nous tenons à remercier Michel Paradis et Brendan Gillon pour leur précieuse aide dans l'organisation du colloque. De même,

nous remercions Hélène Cajolet-Laganière et Louis Mercier qui, en plus de leur participation à l'organisation du colloque, ont aussi collaboré à l'édition de la section sur les lexiques spécialisés. Finalement, nous sommes aussi reconnaissants envers tous les auteurs, tant pour leur contribution originale que pour leur collaboration au cours de toutes les étapes qui ont mené à la publication de ce livre.

Yvan Rose
Julie Auger

# *Table des matières*

**Préface** ..... i

**1. Représentation mentale du lexique bilingue** ..... 1

*Le lexique bilingue à travers la traduction* ..... 3
Kirsten M. Hummel

*Représentation lexicale et conceptuelle chez les bilingues: deux langues, trois systèmes* ..... 15
Michel Paradis

**2. Sémantique lexicale et pragmatique** ..... 29

*L'effet existentiel* ..... 31
Denis Bouchard

*Fonctions configurationnelles* ..... 47
Anna-Maria Di Sciullo

*Le Schéma sémantique diachronique de Stern* ..... 63
Evelyn Matheson-Styan

*Les catégories sémantiques "événement" et "action"* ..... 77
Anne Rochette

*Les connecteurs pragmatiques comme indices du développement du langage chez les enfants de 5 à 11 ans* ..... 89
Madeleine Saint-Pierre

**3. Aspects sociolinguistiques du lexique** ..... 105

*Quels mots les étudiants doivent-ils connaître à leur entrée à l'université?* ..... 107
Annie Côté et Annette Paquot

*Lexicalisation disparate en milieu minoritaire* ..... 121
Terry Nadasdi

*Deux champs sémantiques chez les anglophones et les francophones de Montréal* ..... 133
Gillian Sankoff

*La notion de "francisme" a-t-elle sa raison d'être?* ..... 147
Sylvie Thiboutot et Pierre Martel

**4. Emprunts** ..... 161

*L'absence de flexion sur les emprunts à l'anglais dans le français cadjin* ..... 163
Sylvie Dubois et David Sankoff

*Analyse du contexte d'utilisation de quelques emprunts en vue de leur désambiguïsation automatique* ..... 177
Inès Escayola, Isabelle Lavigne et Michel Théoret

*Lexique bilingue et créativité: modèle de verbes en langue mixte dans une situation de contact entre l'arménien et le français* .... 189
Sylvia Kasparian

*Amérindianismes, variation formelle et dictionnaires* ..... 201
Louis Mercier et Philippe Lanthier

*Préservation phonémique et seuil de tolérance dans l'adaptation d'emprunts* ..... 215
Carole Paradis

*Centre et périphérie: les segments anglais en français québécois* ..... 229
Carole Paradis et Éliane Lebel

*La prononciation des mots anglais en québécois: le problème particulier des anglophones* ..... 243
Marc Picard

*Phonologie et morphologie de la langue emprunteuse: intégration des emprunts français en kinyarwanda* ..... 253
Yvan Rose

**5. Lexiques spécialisés** ..... 265

*Quelques aspects particuliers du traitement prédictionnairique du vocabulaire général d'orientation technique* ..... 267
Hélène Cajolet-Laganière et Normand Maillet

*Quelques critères de sélection des termes dans les dictionnaires bilingues* ..... 281
Aline Francoeur et René Gemme

*Inventaire du vocabulaire de l'alimentation* ..... 293
Josée Mercier

*L'adaptation d'un dictionnaire d'apprentissage à un public québécois: problèmes et solutions* ..... 305
Catherine Ouimet

*Particularités du vocabulaire sociopolitique québécois* ..... 315
Nadine Vincent

# Section 1

## *Représentation mentale du lexique bilingue*

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 3-14.

# Le lexique bilingue à travers la traduction

*Kirsten M. Hummel*
Université Laval

## 1 Introduction

La représentation mentale du lexique bilingue a été examinée par des chercheurs qui ont fait réaliser à des sujets diverses tâches expérimentales. On aimerait ici faire état de quelques études qui ont employé des tâches de traduction pour éclairer la question de l'organisation du lexique bilingue. Tout en reconnaissant que la traduction de mots présentés en dehors d'un contexte linguistique comporte des difficultés, on peut croire que la traduction constitue un outil particulièrement approprié pour examiner des questions relatives à l'organisation du lexique bilingue. D'une part, la traduction est une activité plus naturelle que d'autres tâches expérimentales utilisées dans ce domaine (p. ex., mémorisation de listes de mots, décision lexicale, etc.). D'autre part, pour beaucoup de personnes qui vivent ou qui travaillent dans un milieu où deux langues sont utilisées, la capacité d'employer un mot ou de reformuler un message dans une autre langue est souvent utilisée.

Cependant, on peut examiner de façon critique quelques aspects des études qui ont porté sur la traduction, notamment la méthodologie et l'interprétation des résultats. On fera donc un petit survol des études comportant des tâches de traduction pour ensuite examiner de plus près l'interprétation des résultats de ces études.

### 1.1 Traduction en tant qu'aptitude "naturelle"

D'abord, la traduction peut être examinée en tant qu'aptitude naturelle. Harris et Sherwood (1978) ont proposé que la traduction serait "naturelle", c'est-à-dire qu'elle reposerait sur un ensemble d'aptitudes linguistiques naturelles. Pour ces chercheurs, tout individu connaissant deux langues serait capable de traduire, à partir du moment où il commence à acquérir une deuxième langue. Pour soutenir cette affirmation, ils soulignent l'âge précoce à laquelle la traduction naturelle débute, la prédominance de la traduction spontanée chez les jeunes enfants, le peu de contact nécessaire avec la L2 avant que des enfants plus âgés commencent à traduire, et le manque de corrélation, chez les enfants, entre l'aptitude à traduire et l'instruction en traduction.

Pour ces chercheurs, la traduction serait un aspect

concomitant du bilinguisme, tout comme la capacité de communiquer est liée au fait de connaître une langue. La traduction ne serait pas tellement une capacité apprise mais plutôt se développerait à partir d'une base naturelle existante, comme dans le cas du développement des capacités linguistiques en L1, quoiqu'il puisse y avoir des différences individuelles dans la compétence acquise.

La traduction est souvent perçue comme une capacité accessible seulement aux personnes ayant une formation spécialisée. L'activité de traduction telle que pratiquée par le traducteur professionnel exige en effet des compétences linguistiques très développées. Toutefois, certaines études semblent indiquer que même des enfants peuvent interpréter et traduire des matériaux qui ne dépassent pas leur capacité de compréhension et leur vocabulaire. Ces résultats vont dans le sens de l'affirmation voulant que la traduction soit une aptitude naturelle.

Par exemple, les études faites par Malakoff et ses collègues (Malakoff, 1992; Malakoff & Hakuta, 1991) sur la capacité de traduire chez des enfants bilingues inscrits dans des classes d'éducation bilingue anglais/espagnol aux États-Unis, ont révélé que ces enfants (entre 9 et 12 ans) avaient beaucoup de facilité à traduire et faisaient peu d'erreurs dans leurs traductions, que ce soit vers leur langue première (dans ce cas, l'anglais) ou vers la langue seconde (l'espagnol). Des résultats similaires ont été obtenus, tant pour des enfants habitués à faire des traductions pour des membres de leur famille que pour des enfants ayant moins de pratique de ce genre.

## 2 L'organisation du lexique bilingue

Étant donné que la traduction semble être une aptitude qui, de façon générale, accompagne le fait de connaître une L2, l'étude du processus de la traduction pourrait être une bonne méthode pour aborder la question de l'organisation du lexique bilingue.

Les questions qui se posent dans ce domaine, et sur lesquelles les chercheurs s'interrogent depuis plusieurs décennies, ont souvent tourné autour du nombre de systèmes d'emmagasinage des mots des deux langues chez les bilingues. Pendant plusieurs années, il y avait principalement deux modèles proposés dans la littérature:

1) Un modèle interdépendant: un seul système de représentation pour les deux langues;
2) Un modèle indépendant: deux systèmes séparés, un pour chaque

langue.

Généralement, on cite à l'appui du modèle interdépendant (un système) l'obtention de résultats équivalents dans des conditions unilingues et bilingues. Pour le modèle indépendant (deux systèmes) on se base sur l'obtention de résultats différents pour des conditions unilingues et bilingues. Ensuite, un troisième modèle qui s'accommode mieux des résultats apparemment contradictoires des études a été proposé (Paradis 1980):

3) Un modèle mixte ou hybride: un système unique, au niveau conceptuel, pour les concepts (ou les unités de sens) qui servent d'appui aux deux langues; toutefois cette information conceptuelle serait organisée différemment dans deux systèmes, au niveau lexical.

À l'appui du modèle hybride, on note que des résultats équivalents ou différents entre les conditions unilingues et bilingues dépendent de plusieurs facteurs, surtout le type de tâche et le type de mots utilisés, la façon dont la L2 a été apprise, et aussi le degré de bilinguisme du sujet.

Durgunoglu et Roediger (1987) ont suggéré que les études qui utilisent des tâches faisant appel aux aspects formels des mots (p.ex., décision lexicale, complétion de mot, etc.) fourniront des résultats correspondant au modèle d'indépendance (deux systèmes), tandis que les études qui utilisent des tâches qui font surtout appel au niveau sémantique (le sens) des mots (p. ex., rappel libre, etc.), auront tendance à donner des résultats correspondant au modèle interdépendant (un seul système). Selon ces chercheurs (Durgunoglu & Roediger 1987: 377): "The issue of whether bilinguals store information in one or two codes seems indeterminable, because the varying retrieval demands of different tasks produce different patterns of results and lead to opposite conclusions."

## 3 Traduction vs identification d'images

### 3.1 Modèles de l'"association de mots" et de la "médiation par concept"

Les chercheurs qui ont spécifiquement examiné la traduction sont relativement peu nombreux. Certains chercheurs qui s'intéressent à la traduction comme outil de recherche différencient de la façon suivante deux niveaux de représentation: un niveau lexical

(où sont emmagasinées les formes des mots) et un niveau conceptuel (où sont emmagasinés les contenus sémantiques des mots). De plus, ces chercheurs prennent pour acquis l'existence de deux lexiques séparés, un pour chaque langue du bilingue. Ils proposent deux modèles possibles pour expliquer ce qui se passe lors de la traduction de mots.

Dans sa forme originale, le modèle de "l'association de mots" a proposé des associations directes entre les représentations lexicales de mots équivalents dans les deux langues du sujet bilingue. Il était proposé que les mots de la L2 sont associés aux mots de la L1 et que c'est à travers les mots de la L1 que les mots de la L2 ont accès aux concepts.

Le modèle de la "médiation par concept" dans sa forme initiale a proposé que le seul lien entre les deux lexiques serait le système conceptuel, auquel les images ont aussi accès. Dans ce modèle, les mots de la L2 ont accès directement aux concepts tout comme les images.

Les études sur ces modèles comparent généralement une tâche de traduction avec une tâche d'identification d'images. Selon Snodgrass (1993), il y a deux raisons principales pour effectuer une telle comparaison:

1) une mesure du temps pour traduire d'une langue à une autre devrait fournir une façon relativement directe d'examiner les étapes relatives à l'accès au lexique bilingue.

2) la comparaison de tâches de traduction avec des tâches d'identification d'images se fait parce que ces dernières ont été l'objet de plusieurs recherches qui font ressortir les ressemblances entre ces deux processus:

a--Pour les deux tâches, les sujets doivent émettre une réponse vocale.
b--Dans les deux cas, la variable dépendante qui nous intéresse est la vitesse de réponse.
c--Et dans les deux situations, le sujet doit utiliser des connaissances sémantiques pour décoder la représentation physique du stimulus à partir de sa forme nominale vers la langue d'énoncé.

Les chercheurs ont donc tendance à se concentrer sur une comparaison entre l'identification des images en L2 et la traduction d'un mot de la L1 en L2. La plupart des chercheurs s'entendent pour

dire que l'identification des images doit passer par une médiation par concept, c'est-à-dire que pour nommer une image, il est d'abord nécessaire d'accéder au sens de l'image avant d'accéder à son nom. Donc, dans un sens, identifier une image serait une sorte de tâche de traduction dans laquelle l'image représente une langue d'entrée (d'input) et la L1 ou la L2 seraient les langues d'émission (d'output).

De plus, les chercheurs sont généralement d'accord sur le fait que, chez les bilingues, la traduction doit également passer par une médiation par concepts, c'est-à-dire que pour traduire de la L1 à la L2, ou vice versa, le locuteur doit d'abord accéder au sens du mot stimulus avant d'accéder au mot réponse.

Toutefois, on trouve quelques points de vue différents concernant cette dernière hypothèse dans la littérature. Entre d'autres, il est proposé que:

1--pour des bilingues débutants, la traduction de la L1 vers la L2 se ferait par association de mots directe entre les deux langues.
2--la traduction vers la L1, au lieu de passer par des concepts, se ferait par une association de mots.

### 3.2 Apprenants débutants vs avancés

On fait le raisonnement que si les apprenants débutants utilisaient des associations de mots, leur temps de traduction de la L1 vers la L2 devrait être plus rapide que le temps pour nommer des images. Cette supposition est basée sur l'hypothèse que le nombre d'étapes entre le stimulus et la réponse détermine le temps de réaction. Puisqu'il n'y aurait pas d'étape d'accès sémantique si la traduction suivait le modèle d'association de mots, alors que cette étape existe pour nommer une image, la traduction devrait être plus rapide.

Ainsi, le modèle de l'association de mots prévoit que la traduction en L2 prendra moins de temps que de nommer une image en L2 pour la même raison que la lecture d'un mot en L1 prend moins de temps que de nommer une image en L1, puisque nommer une image doit passer par une étape supplémentaire de récupération de concept.

#### 3.2.1 Potter, So, Von Eckardt, & Feldman, 1984

Une des premières études à examiner ces modèles en utilisant des tâches de traduction a été celle de Potter et al. (1984). Ils ont fait

la comparaison entre un groupe de bilingues chinois/anglais qui avaient une bonne compétence en L2 et un groupe de bilingues anglais/français, moins à l'aise en L2, qu'on pourrait qualifier de "débutants".

Les sujets étaient des "apprenants compétents" (quoiqu'aucun test de compétence ne soit mentionné): n= 24, L1= chinois, L2= anglais. Éduqués en pays de langue chinoise, ils avaient appris l'anglais L2 à l'école et avaient habité et étudié en pays anglophone un an ou plus.

Les sujets étaient exposés aux stimuli par tachistoscope. Les stimuli étaient, d'une part, des dessins d'objets et, d'autre part, les noms de ces objets écrits en chinois et en anglais. Chaque item était présenté pendant 250 msec et les sujets devaient nommer les objets ou traduire les noms des objets. La moitié du temps la langue de réponse était le chinois, pour l'autre moitié, c'était l'anglais.

Les résultats qui nous intéressent concernent le temps nécessaire pour nommer une image en L2 (anglais), et traduire un mot L1 (chinois) en L2 (anglais), par opposition au temps nécessaire pour nommer une image ou un mot de la L1 en L1. Nommer les images en L1 prenait plus de temps que de nommer le mot de la L1 en L1, tel que prédit par le modèle "médiation par concepts." Également conséquentes avec le modèle de la "médiation par concept", la traduction de mots en L2 et l'identification des images étaient faites sensiblement à la même vitesse.

Dans la deuxième partie de cette recherche, Potter et al. ont voulu savoir si des apprenants moins avancés en L2 seraient plus dépendants des associations entre les mots des deux langues, c'est-à-dire qu'ils suivraient davantage le modèle de l'association de mots. Les sujets "apprenants débutants" étaient 28 élèves de niveau secondaire aux États-Unis, de langue maternelle anglaise, ayant étudié le français (L2) pendant 2 ou 3 ans.

Les résultats n'ont pas validé le modèle de l'association de mots et les résultats ressemblaient beaucoup à ceux des sujets plus avancés de la première étude.

Donc, dans cette étude, Potter et al. ont trouvé des éléments à l'appui uniquement du modèle de médiation par concept puisque les temps pour traduire et pour nommer des images en L2 étaient similaires. À l'encontre de l'hypothèse voulant que les sujets moins compétents en L2 passeraient par une association directe d'un mot à l'autre, aucune différence n'a été trouvée entre ces sujets et ceux qui étaient plus avancés en L2. Potter et ses collègues ont donc conclu qu'une médiation par les concepts est possible très tôt dans

l'apprentissage d'une L2.

### 3.2.2 Chen & Leung, 1989

Une autre étude (Chen & Leung, 1989) sur la même question a consisté à comparer encore une fois une tâche de traduction et une tâche d'identification d'images. Les sujets avaient tous la même langue maternelle, le chinois, mais la L2 était différente et il y avait deux groupes de compétences différentes en L2.

Les sujets qui nous intéressent étaient 2 groupes de locuteurs de langue chinoise, chaque groupe comptant 18 sujets.

1) Les sujets "compétents" étaient des étudiants à l'Université de Hong Kong avec plus de 12 ans de cours d'anglais

2) Les sujets "débutants" étaient des adultes étudiant à l'Université de Hong Kong qui avaient étudié le français pendant 2 à 3 ans

La méthodologie suivie était très similaire à celle de Potter et al. (1984). Les sujets devaient nommer des dessins en L1 ou L2 ou traduire des mots chinois en L2 (anglais ou français selon le groupe).

Les résultats indiquent que le groupe ayant le moins d'expérience en L2 confirme les hypothèses du modèle de l'association de mots: ils ont effectué la traduction de la L1 vers la L2 plus rapidement qu'ils ont identifié les images en L2, ce qui laisse supposer qu'ils ont associé les représentations lexicales sans passer par une étape conceptuelle. Les sujets les plus compétents en L2 ont eu des résultats comparables à ceux de Potter et al., c'est-à-dire que les temps de traduction et d'identification d'images étaient similaires, ce qui indique un passage par un niveau conceptuel dans les deux cas.

### 3.2.3 Critique

Chen et Leung ont estimé que les sujets moins avancés de l'étude de Potter et al. avaient déjà dépassé l'étape où le modèle de l'association de mots prédomine, alors que leurs sujets moins avancés étaient toujours à l'intérieur de cette étape.

Pourtant il convient d'examiner certains aspects de ces études. Notamment, on peut prêter attention au choix des sujets. Par exemple, dans l'étude de Potter et al., les deux groupes comparés sont constitués de bilingues de deux paires de langues très différentes ayant des expériences linguistiques très différentes:

--L1 (le chinois (une langue idéographique) vs l'anglais),
--L2 (l'anglais vs le français),
--âge d'acquisition de la L2 (école primaire vs secondaire),
--âge au test (université vs école secondaire)

Néanmoins, les auteurs ont tenté de dégager des conclusions en comparant deux populations aussi différentes. On peut croire que le traitement cognitif associé à des langues aussi peu semblables sur le plan typologique pourrait rendre la comparaison difficile.

On peut également faire remarquer que, dans l'étude de Chen et Leung, quoique tous les sujets aient la même langue maternelle, à savoir le chinois, la L2 est différente pour les sujets jugés avancés (L2= anglais) et les sujets considérés débutants (L2=français). Ce ne sont donc pas des populations idéales pour nous permettre de tirer des conclusions fiables.

### 3.3 L'asymétrie de traduction (L2 vers L1, L1 vers L2)

Un résultat qu'on obtient systématiquement est que la traduction vers la L1 est plus rapide que celle effectuée vers la L2. Pour Kroll et Sholl (1992), ce résultat laisse perplexe et les auteurs ont supposé que cette asymétrie indique que la traduction de la L2 vers la L1 se réaliserait par la voie d'une association directe de mots (puisque plus rapide), tandis que la traduction de la L1 vers la L2 se ferait principalement par concepts (puisque moins rapide).

Donc, dans le modèle révisé proposé par Kroll et Stewart (1992), les liens lexicaux sont plus forts dans le sens de la L2 vers la L1 que vice versa, mais les liens conceptuels pour la L1 sont plus forts que ceux qui existent pour la L2.

Dans ce modèle révisé, on explique l'asymétrie en traduction par le fait que la présentation d'un mot en L1 peut davantage activer un concept que son équivalent lexical en L2. Pourtant, la présentation d'un mot en L2 est plus susceptible d'activer le mot en L1 qui est son équivalent, que le concept. Dans ce modèle il est plus probable que la traduction vers la L2 implique un traitement conceptuel et donc prend plus de temps que la traduction vers la L1.

Ces chercheurs affirment que ce modèle est basé sur l'hypothèse que la traduction vers la L2 devrait être influencée par des variables d'ordre sémantique ou conceptuel, tandis que la traduction vers la L1 ne serait pas sensible au contexte sémantique.

### 3.3.1 Kroll & Stewart, 1992

Pour ces chercheures, puisque le lien de la L1 vers la L2 passe par la médiation des concepts, une manipulation qui affecte l'accès conceptuel ou sémantique devrait affecter la traduction vers la L2, mais pas l'opposé. Donc, elles ont organisé des mots à être traduits dans des listes, par catégories sémantiques, versus des listes où les mots étaient présentés de façon aléatoire.

Les sujets bilingues (hollandais/anglais) avaient à identifier des mots et à traduire les mots présentés dans les listes. On a obtenu comme résultat que seulement la traduction vers la L2 était affectée par le contexte sémantique; pour la traduction vers la L1, il n'y avait pas d'effet de la variable de catégorisation sémantique. La direction de cet effet en était une d'inhibition, i.e., la traduction vers la L2 était ralentie lorsque les mots étaient présentés par catégories.

Kroll et Stewart concluent que ces résultats confirment l'hypothèse voulant qu'une traduction de la L1 à la L2 exige un traitement conceptuel. De plus, le fait qu'une catégorisation sémantique n'ait pas eu d'effet sur la traduction vers la L1 appuierait l'hypothèse que la route de la L2 à la L1 passe par des connections lexicales.

### 3.3.2 Critique

On peut cependant trouver des explications plus vraisemblables. Il y a plusieurs arguments pour expliquer l'asymétrie de traduction. On peut notamment souligner que lorsqu'on n'est pas en situation de "bilinguisme équilibré" le fait de traduire vers sa langue dominante soit plus rapide que de traduire vers la langue seconde est loin d'être "inattendu", pas plus que le fait de s'exprimer dans sa langue maternelle se fasse plus facilement qu'en L2. Dans les deux cas, il s'agit d'un accès plus rapide à un lexique bien connu par rapport à celui d'une langue moins bien connue, tout comme on constate des différences entre des tâches impliquant des mots à forte ou à faible occurrence: en général, plus une tâche (ou des mots) est familière, plus son exécution est rapide. Donc, l'hypothèse voulant que la traduction vers la L1 se fasse par des liens lexicaux seulement, sans passer par des liens conceptuels, n'est pas convaincante. On pourrait même dire qu'il est contre-intuitif de suggérer que la traduction peut court-circuiter le niveau conceptuel, que ce soit vers la L1 ou vers la L2.

De plus, si l'on considère les résultats des études dont on a

discuté au début de cet article ainsi que les résultats de ces dernières études, on constate qu'on finit par avoir le même modèle proposé pour des débutants en L2 traduisant vers la L2 (la langue faible) et pour des apprenants relativement avancés qui traduisent vers la L1 (la langue dominante); dans les deux cas, avec le modèle de l'association de mots proposé, les liens lexicaux devaient prédominer. On peut croire que ces deux extrêmes (apprenants débutants et bilingues avancés traduisant dans leur langue dominante) sont difficilement réconciliables à l'intérieur d'un seul et même modèle.

## 4 Conclusions

Pour résumer, deux modèles ont été proposés dans la littérature dont on a discuté en ce qui à trait à la traduction. Généralement, certains chercheurs ont proposé que, tout comme l'identification d'une image doit obligatoirement passer par une étape conceptuelle, traduire dans une L2 passe aussi par une médiation de concept. Néanmoins, plusieurs chercheurs ont proposé que les apprenants débutants passent par une association directe de mots entre les items des deux langues. Il faut souligner, pourtant, que cette affirmation n'est que faiblement appuyée par les résultats obtenus. Il est vrai que dans l'étude de Chen et Leung (1989) les débutants ont eu des performances plus rapides pour la traduction que pour l'identification d'images, mais comme on a fait remarquer, les différences entre les groupes de sujets avancés et débutants (L2 anglais vs L2 français) mettent en question la fiabilité de ces résultats.

Deuxièmement, en ce qui à trait à l'asymétrie de la traduction, le même modèle de l'association de mots est proposé dans la littérature. La traduction vers la L1 se ferait selon le modèle de l'association de mots, sans nécessairement passer par le niveau conceptuel. Mais, comme nous l'avons proposé, la rapidité de traduction vers la langue dominante peut très bien s'expliquer par d'autres facteurs, entre autres, la plus grande familiarité avec les mots de sa langue maternelle et conséquemment une plus grande facilité d'accès au lexique correspondant.

Il y a d'autres aspects qui limitent la portée des recherches discutées qu'on n'a pas eu le temps d'aborder, par exemple, le fait que seulement des mots concrets, ayant des équivalents de traduction évidents, ont été utilisés dans ces études pour qu'ils soient facilement asssociés aux images.

On peut conclure en soulignant que, pour être réellement fidèle à ce qui se passe lors du traitement cognitif qui accompagne la

traduction, les mots devraient être situés dans un contexte grammatical. Il y a peu de mots qui ont un seul et unique équivalent de traduction. Le traducteur dépend du contexte pour bien choisir le mot approprié. Sinon, on peut obtenir des anomalies du genre "ventilateurs de sport" (pour "sports fans"), mauvaise traduction parue récemment dans une annonce publiée dans la région de Québec, qui n'ont rien à voir avec une traduction fidèle. Les tâches de traduction utilisées dans les études examinées dans cet exposé sont en somme limitées dans leur capacité de nous donner une perception exacte de la représentation du lexique bilingue.

## Bibliographie

Chen, Hsuan-Chih & Yuen-Sum Leung. 1989. "Patterns of lexical processing in a nonnative language." *Journal of Experimental Psychology*, 15: 316-325.

Durgunoglu, Aydin & Henry L. Roediger III. 1987. "Test differences in accessing bilingual memory." *Journal of Memory and Language*, 26: 377-391.

Harris, Brian & Bianca Sherwood. 1978. "Translating as an innate skill." In D. Gerver & H.W. Sinaiko (réd.). *Language interpretation and communication*. New York: Plenum Press, pp. 155-170.

Kroll, Judith F. & Alexandra Sholl. 1992. "Lexical and conceptual memory in fluent and nonfluent bilinguals." In R.J. Harris (réd.). *Cognitive processing in bilinguals*. Amsterdam: North-Holland, pp. 191-204.

Kroll, Judith F. & E. Stewart. 1992. "Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations." Manuscrit inédit. Mount Holyoke College, South Hadley, MA.

Malakoff, Marguerite E. 1992. "Translation ability: A natural bilingual and metalinguistic skill." In R.J. Harris (réd.). *Cognitive processing in bilinguals*. Amsterdam: North-Holland, pp. 515-529.

Malakoff, Marguerite E. & Kenji Hakuta. 1991. "Translation skill and metalinguistic awareness in bilinguals." In E. Bialystok (réd.). *Language processing and language awareness in bilingual children*. NY: Oxford University Press, pp. 141-166.

Paradis, Michel. 1980. "Language and thought in bilinguals." In

H.J. Izzo & W.E. McCormack (réd.). *The sixth LACUS forum*. Columbia, SC: Hornbeam Press, pp. 420-431.

Potter, Mary C., Kwok-Fai So, Barbara Von Eckardt, & Laurie B. Feldman, 1984. "Lexical and conceptual representation in beginning and proficient bilinguals." *Journal of Verbal Learning and Verbal Behavior*, 23: 23-38.

Snodgrass, Joan Gay. 1993. "Translating versus picture naming: similarities and differences." In R. Schreuder et B. Weltens (réd.). *The bilingual lexicon*. Philadelphie: John Benjamins, pp. 83-114

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 15-27.

# Représentation lexicale et conceptuelle chez les bilingues: deux langues, trois systèmes

*Michel Paradis*
Université McGill

Au cours des années 60 et 70, les psychologues se posaient la question de savoir si le sujet bilingue possédait deux réservoirs mnésiques, soit un pour chaque langue, ou un seul réservoir commun aux deux langues. Les expériences psycholinguistiques de l'époque sur ce que l'on appelait "la mémoire linguistique bilingue" suivaient toutes un raisonnement de base semblable: Les mêmes tâches étaient présentées sous deux conditions, une condition unilingue et une condition bilingue; que ce soit le rappel libre de listes de mots, le rappel de listes de mots arbitrairement appareillés, les associations de mots, ou l'évitement d'un choc électrique. Si le comportement des sujets dans la condition bilingue était identique à celui de la condition unilingue, les résultats étaient interprétés comme confirmant l'hypothèse d'un réservoir mnésique commun; si les résultats étaient différents d'une condition à l'autre, ils étaient interprétés comme confirmant l'hypothèse de deux réservoirs distincts.

Inutile de dire que les résultats rapportés dans la littérature psycholinguistique étaient contradictoires, certains étant interprétés comme soutenant l'hypothèse de deux réservoirs, d'autres, comme soutenant celle d'un seul. Or, comme le fit remarquer Kolers en 1968 (Fig. 1), les résultats de ses propres recherches ne confirmaient ni l'une ni l'autre hypothèse. Ses sujets avaient des résultats trop différents pour conclure à un réservoir commun, mais pas suffisamment pour conclure à deux réservoirs séparés. Les mots et leur traduction se comportaient comme les synonymes d'une même langue.

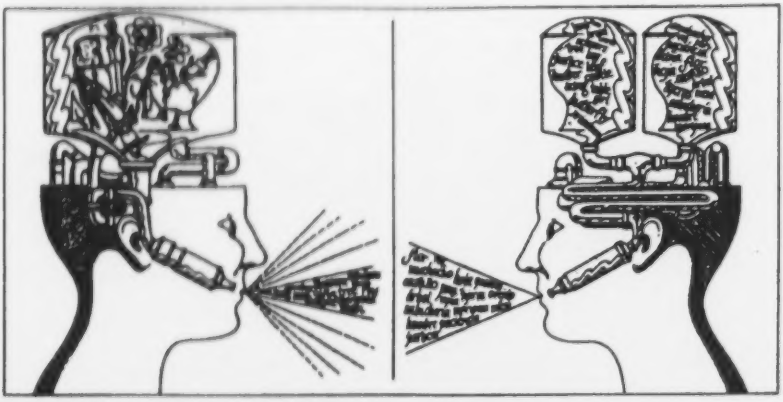
De fait, toutes ces expériences souffraient d'un nombre d'inconvénients qui les rendaient ininterprétables:

(1) Elles ne précisaient pas ce qu'elles entendaient par "mémoire linguistique bilingue" (soit le système de la langue, soit l'information acquise verbale-ment).
(2) Elles ne tenaient pas compte des différents types de bilinguisme de leurs sujets (c'est-à-dire de l'organisation de la signification des mots, selon qu'ils contiennent une interférence unidirectionnelle, que l'on appellera organisation subordonnée du lexique; une interférence bidirectionnelle, ce que Weinreich

(1968) a appelé sytème amalgamé; ou sans interférence, ce qu'il a appelé systèmes coexistants).

(3) Elles ne distinguaient pas entre mémoire à court terme pour les mots (c.à.d. sens + forme) et la rétention en mémoire à long terme du message (Fig. 2).

(4) Et surtout, elles ne faisaient pas de distinction entre les unités de sens (qui font partie du système lexical) et les représentations conceptuelles (qui sont indépendantes du langage, bien qu'elles interagissent avec lui).



Two hypotheses about the way a bilingual person handles information are represented by two arrangements of tanks. One hypothesis (left) is that all his information is stored centrally, or in one tank, and that he has access to it equally with both languages, which are represented by the various taps. The other (right) is that his information is stored in linguistically associated ways, or in separate tanks. Experiments by the author indicated that the actual situation of a bilingual person combines parts of both hypotheses. (Kolers, 1968)

**Figure 1.** Un ou deux réservoirs mnésiques?

Toutes ces expériences n'ont porté que sur la phase 2 (Fig. 2): la rétention en mémoire à court terme de mots en dehors de tout contexte significatif. Mais même lorsque l'on ne s'occupe que des mots, comme c'est explicitement le cas dans les études plus récentes de la fin des années 80 et début 90, il est important de distinguer entre un réservoir mnésique conceptuel indépendant du langage (et donc des langues) et une mémoire sémantique propre à la langue, comme le démontrent le cas d'aphasie paroxistique décrit par

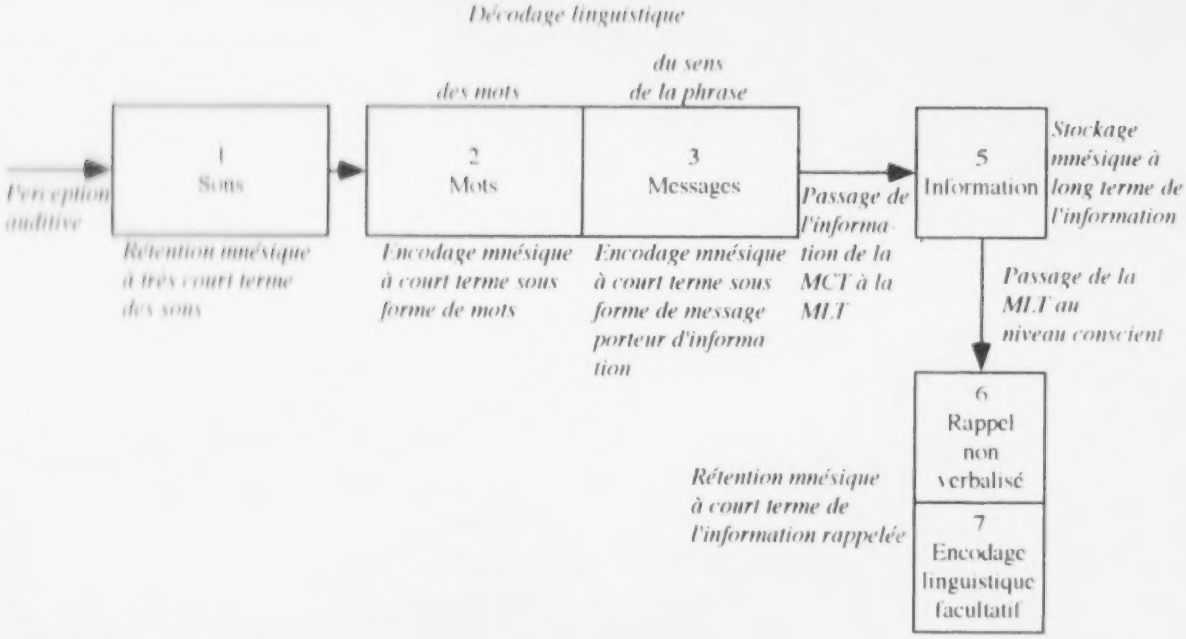



**Figure 2.** Décodage linguistique et encodage mnésique de l'information transmise verbalement

Lecours et Joanette (1980). En effet, il existe une distinction très nette entre le système concepuel, qui demeure relativement intact chez les aphasi-ques, et le système sémantique lexical, qui peut être sélectivement affecté.

Les bilingues, dans la mesure où ils n'ont pas de difficulté de communication dans chacune de leurs langues, possèdent nécessairement deux systèmes linguistiques, de même qu'un seul système conceptuel, dans lequel les représentations sont organisées différemment selon qu'elles correspondent aux contraintes sémantiques lexicales de l'une ou de l'autre langue. La solution que j'ai proposée en 1978 est illustrée dans la Figure 3.

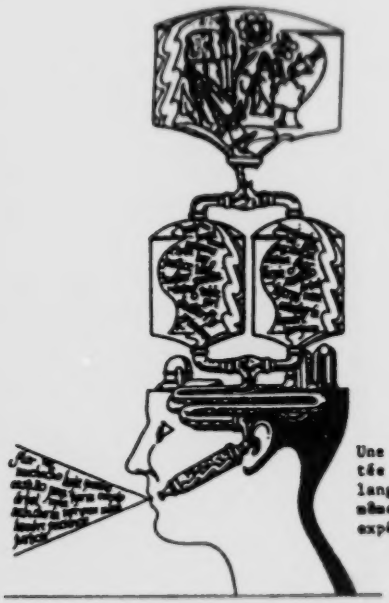



**Figure 3**. L'hypothèse des trois réservoirs

La plupart des modèles contemporains de la représentation mentale du lexique distinguent effectivement entre deux niveaux de représen-tation: le niveau lexical (qui contient les formes lexicales) et le niveau conceptuel (qui contient le sens).

Toutefois, l'hypothèse des trois systèmes (Paradis, 1978, 1980,

1986) considère le sens des mots aussi bien que la forme lexicale (la forme phonologique et les propriétés syntaxiques) comme faisant partie de la représentation lexicale, aux côtés d'un troisième système de représentations, conceptuel celui-là, indépendant des langues. Deux lexiques sont ainsi postulés, nécessairement différents, tel que démontré par les nombreux cas d'atteinte sélective et de récupération non-parallèle des langues chez les aphasiques bilingues (Paradis, 1989). On postule donc deux lexiques, un pour chaque langue, plus un système conceptuel unique, quel que soit le nombre de langues.

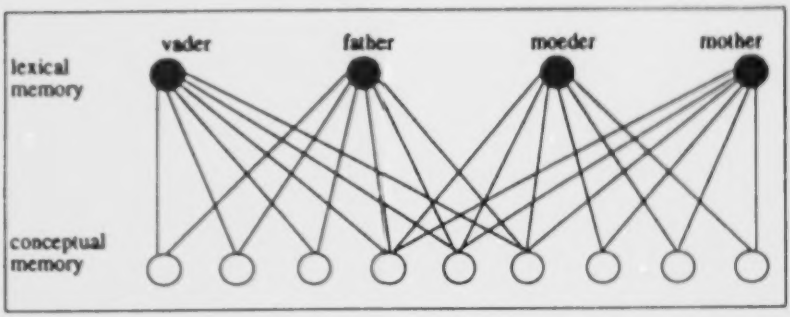
Dans ce modèle, le mot est représenté comme une matrice multidimensionnelle de traits phonologiques, syntaxiques, morphologiques et sémantiques. La représentation d'un item lexical comprend donc non seulement une forme lexicale, mais aussi des contraintes sémantiques s'appliquant à son sens (c'est-à-dire à ce qu'il peut désigner). Pour autant qu'un mot et sa traduction ne contiennent pas toutes les mêmes contraintes sémantiques (sans parler des propriétés phonologiques, morphosyntaxiques et graphémiques), le lexique doit nécessairement être spécifique à chaque langue, c'est-à-dire avoir une entrée indépendante pour chaque mot et pour sa traduction, y compris son sens. Chaque item lexical détermine alors quels traits conceptuels seront activés lorsque le mot est utilisé. Les traits conceptuels font partie du système de représentations mentales non linguistiques qui sont phylogénétiquement et ontogénétiquement antérieures au développement du lexique et qui continuent à se développer avec l'expérience (y compris l'expérience linguistique). Chaque mot active un ensemble correspondant de traits sémantiques de manière à obtenir une représentation mentale de son référent (objet, qualité, événement, etc.). La représentation mentale ainsi activée par un mot différera de la représentation conceptuelle de sa traduction dans la mesure où leurs contraintes sémantiques lexicales diffèrent d'une langue à l'autre. La représentation du mot français "cheveu" correspond à un poil qui vient de sur la tête, pas de sous le nez, sous le menton, ou sous les bras, et pas de sur la tête d'un animal; la représentation qui correspond au mot anglais "hair" réfère à un poil humain ou animal provenant de n'importe quelle partie du corps, y compris sur la tête. *A tiger is a cat*, mais un tigre n'est pas un chat. "*A ball*" ne correspond pas aux mêmes référents qu'une balle, pas plus que "une chaise" à ceux de "*a chair* ". (Une balle doit être suffisamment petite pour tenir dans la main, sinon, c'est un ballon; une chaise ne peut avoir d'accoudoir, sinon c'est un fauteuil. En anglais, quelle que soit la taille, c'est toujours "a ball"; avec ou sans accoudoir, c'est toujours "a chair".)

Une fois que les concepts ont été acquis, même verbalement, ils sont représentés indépendamment du langage et peuvent être activés sans recours au langage. Ce troisième système demeure disponible au patient aphasique (Hécaen, 1968; Lecours & Joanette, 1980). Les données en provenance de l'aphasie globale (Gardner et al., 1976), de la jargonaphasie (Beauvois & Dérouesné, 1976), du syndrome de Down (Rondal, 1994) et du syndrome de Williams (Rondal & Comblain, 1995), suggèrent que les contraintes sémantiques lexicales existent indépendamment des représentations conceptuelles non linguistiques. Les contraintes sémantiques de chaque mot imposent des restrictions aux représentations conceptuelles correspondantes et donc déterminent quelle représentation sera activée. Inversement, lors de l'encodage, les représentations mentales doivent s'accommoder des contraintes sémantiques idiosyncratiques impo-sées par chaque langue. Bien que les contraintes sémantiques lexicales spécifient quels traits conceptuels sont activés, les spécifications lexicales elles-mêmes demeurent partie intégrante du système de la langue (c.à.d. spécifiques au français, au malgache ou au japonais).

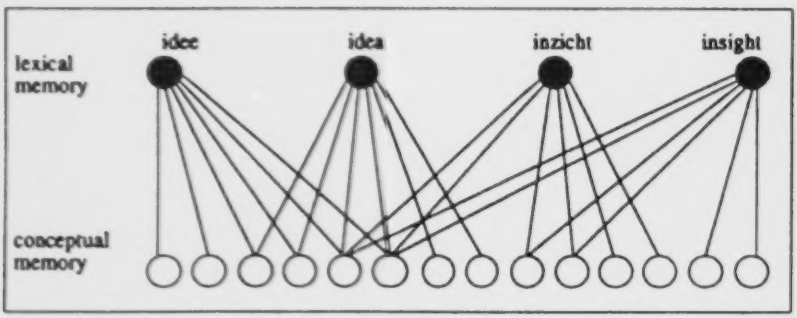
Un nombre croissant de chercheurs adoptent explicitement ou implicitement le modèle des trois systèmes (Grosjean, 1982; Perecman, 1984; Grainger, 1987; Klosty Beaujour, 1989; Wray, 1992; Meyers-Scotton, 1993; Hagège, 1996) ou du moins distinguent entre un niveau sémantique et un niveau conceptuel (Levelt, 1989; de Groot, 1982; Green, 1993, de Bot et al., 1995). Toutefois, la plupart du temps, les représentations conceptuelles sont présumées communes aux deux langues. Si c'était le cas, les bilingues se trouveraient dans la position peu enviable de ne pouvoir communiquer efficacement avec des locuteurs de leur deuxième langue (s'ils avaient un système subordonné) ou avec les locuteurs ni d'une langue ni de l'autre (s'ils avaient un système amalgamé).

Le modèle qui se rapproche le plus de l'hypothèse des trois systèmes est celui de de Groot (1992), dans le cas de sa figure 4b reproduite dans la Figure 4, qui est la situation qu'elle semble toutefois n'attribuer qu'aux mots abstraits. En effet, le sens des mots allemands et anglais représentés en mémoire lexicale correspond à un ensemble de traits en mémoire conceptuelle, dont seulement certains sont partagés par un mot et sa traduction (idee, idea), tout comme certains traits sont partagés par un mot et son synonyme dans la même langue (idea, insight; inzicht, idee).

Malgré tout, de Groot continue à s'exprimer comme s'il y avait deux types d'organisation conceptuelle, l'une spécifique à la langue, pour les mots abstraits, l'autre commune aux deux, pour les mots concrets. Cela revient à confondre les résultats obtenus aux expériences et la raison pour laquelle ils sont obtenus. En effet, il n'y a pas de raison de postuler deux types de représentations qui correspondraient respectivement aux mots concrets et abstraits (ou toute autre catégorie dichotomique).




4a




4b

**Figure 4.** Représentation conceptuelle (4a) mots concrets, (4b) mots abstraits d'après de Groot (1982)

Tout d'abord, il ne faudrait pas minimiser les différences entre les mots concrets et leur traduction la plus proche. Non seulement les connotations et les sens figuratifs mais aussi le sens dénotatif diffère plus qu'on a tendance à le croire. Rares sont les mots (s'il y en a) dont la traduction désigne exactement les mêmes référents. Bien sûr, comme le fait remarquer de Groot (1992, p. 404), *a chair* de même qu'une chaise sont faites pour s'asseoir dessus, toutefois il y a des tas de *chairs* qui ne sont pas des chaises. Si je demande à quelqu'un de m'apporter toutes les chaises qui sont dans le salon, la personne ne me rapportera pas les mêmes objets que si je demande qu'on m'apporte *all the chairs*, car dans ce cas-là on m'apporterait aussi les fauteuils (Fig. 5). Et les pommes, dans toutes les langues, sont pour être mangées, dit-elle; oui, mais là où le français distingue entre les pommes et les poires, certaines langues, comme le japonnais, décomposent la réalité différemment.

Ensuite il n'y a pas de raison de postuler deux types différents de représentations, l'un pour les mots abstraits, représentés dans un système spécifique à la langue, l'autre pour les mots concrets, commun aux deux langues. Selon de Groot, les données expérimentales suggèrent que le type de représentation dépend vraisemblablement du type de mot (abstrait ou concret, émotif ou neutre, apparenté ou non), mais, et je cite (du moins je traduis), "le degré de chevauchement du sens d'un mot particulier et de son équivalent le plus proche dans l'autre langue est le facteur critique qui détermine la manière dont les deux mots sont emmagasinés en mémoire." (p. 404). Si le degré de chevauchement du sens d'un mot et de sa traduction est le facteur critique, pourquoi maintenir une dichotomie? Les différences d'un mot à l'autre sont sur un continuum, puisque déterminées par le nombre de traits sémantiques partagés, un phénomène tout à fait aléatoire et qui, pour chaque item lexical, change à l'intérieur de chaque paire de langues. De Groot a tous les éléments en main, et cependant elle propose un modèle en termes de "deux types de représentation", un système spécifique à la langue (pour les items les plus différents), et un système commun (pour les items les plus semblables).

La différence entre ce que de Groot semble proposer et ce que je propose, c'est que pour elle il y a deux types de représentations, et que pour moi, il n'y a qu'un système conceptuel, dont certains traits sont activés selon les contraintes sémantiques lexicales propres aux mots de chaque langue. Le nombre de traits partagés varie d'un mot

32

**Figure 5.** Annonce publicitaire parue dans un journal montréalais de langue anglaise.
Il n'y a aucun doute que les anglophones appellent "chair" ce que les francophones appelleraient "fauteuil".

à l'autre dans chaque paire de langues, selon un continuum et non une dichotomie, qu'ils soient concrets ou abstraits (même s'il n'est pas invraisemblable qu'il y ait plus de différences entre les mots abstraits qu'entre les mots concrets—ce qui les distingue n'est pas leur nature concrète ou abstraite, mais dans les deux cas, leur degré de chevau-chement). Dans la mesure où un mot et sa traduction partagent des traits sémantiques, ils activeront les mêmes traits conceptuels. Dans la mesure où ils possèdent des traits différents, les représen-tations conceptuelles différeront. Cela suffit à expliquer les différences de comportement de divers mots au cours de tâches expérimentales comme celle de l'amorçage interlinguistique, par exemple. Le résultat des expériences ne fait que refléter cette différence de degré de chevauchement.

Pour conclure, tout modèle de la représentation mentale du lexique chez les bilingues doit nécessairement rendre compte des propriétés sémantiques lexicales d'une part et des propriétés conceptuelles d'autre part. Les représentations conceptuelles correspondant aux mots et celles correspondant à leur traduction dans les différentes langues ne sauraient être identiques dans la mesure où l'individu évite les malentendus.

Chaque langue organise les représentations mentales en accord avec ses propres contraintes sémantiques.

Les mots équivalents dans une autre langue sont comme des synonymes dans la même langue en ce que leur sens n'est jamais identique. Les représentations mentales sont évoquées selon le sens correspondant à chaque mot. Elles différeront donc d'une langue à l'autre. Les représentations mentales n'en demeurent pas moins isolables du lexique. (comme le démontre la pathologie).

Dans la mesure où les locuteurs parlent leurs deux langues sans interférence (Paradis, 1993), il n'y a donc pas une représentation sémantique commune aux mots et à leurs équivalents en traduction, mais un seul système cognitif, indépendant du langage, et donc des langues, dans lequel différentes représentations mentales correspondent au sens d'un mot d'une langue et à celui de sa traduction dans l'autre. Le sens dénotatif des mots français concrets "chaise", "balle", "cheveu", "chat", et autres, n'évoque pas la même représentation mentale que leur équivalent respectif le plus proche, *chair, ball, hair, cat*.

## Références

Beauvois, F. & J. Dérouesné. 1976. Un modèle de fonctionnement de la mémoire verbale, ses implications en pathologie cérébrale. *La mémoire sémantique* (numéro spécial annuel du *Bulletin de psychologie*) pp. 166-173.

de Bot, K., Cox, A., Ralston, S. Schaufeli, A., & B Weltens. 1995. Lexical processing in bilinguals. *Second Language Research*, 11, 1-19.

De Groot, A.M.B. 1992. Bilingual lexical representation: A closer look at conceptual representations. In R. Frost & L. Katz (Eds.), Orthography, phonology, morphology, and meaning. Amsterdam: Elsevier, pp. 389-412.

Grainger, J. 1987. L'accès au lexique bilingue: Vers une nouvelle orientation de recherche. *L'Année Psychologique*, 87, 553-566.

Green, D.W. 1993. Toward a model of L2 comprehension and production. In R. Schreuder & B. Weltens (eds.), *The bilingual lexicon*. Amsterdam: John Benjamins, pp.249-277

Grosjean, F. 1982. *Life with two languages*. Cambridge, MA.: Harvard University Press.

Hagège, C. 1996. *L'enfant aux deux langues*. Paris: Odile Jacob.

Hécaen, H. 1968. L'aphasie. Dans A. Martinet (Ed.), *Le langage*. Paris: Gallimard, pp. 390-414.

Klosty Beaujour, E. 1989. *Alien Tongues*. Ithaca: Cornell University Press.

Kolers, P. 1968. Bilingualism and information processing. *Scientific American*, 218 (March) 78-86.

Lecours A.R. & Y Joanette. 1980. Linguistic and other psychological aspects of paroxismal aphasia. *Brain and Language*, 10, 1-23.

Levelt, W.J.M. 1989. *Speaking. From intention to articulation*. Cambridge, MA: MIT Press.

Luria A.R. 1974. Basic problems of neurolinguistics. In T.A. Sebeok (ed.) *Current trends in linguistics*, Vol. 12. The Hague: Mouton, pp. 2561-2593.

Meyers-Scotton, C. 1993. *Duelling languages: Grammatical structure in code-switching*. Oxford: Clarendon Press.

Paradis, M. 1978. Bilingual linguistic memory: Neurolinguistic considerations. Communication présentée à la Linguistic Society of America, Boston, 28 décembre.

Paradis, M. 1980. Language and thought in bilinguals. In W.C. McCormack & H.J. Izzo (eds.), *The sixth LACUS forum*. Columbia, SC: Hornbeam Press, pp. 420-431.

Paradis, M. 1986. Bilinguisme. Dans J. Rondal et J.-P. Thibaut (eds.), *Problè-mes de psycholinguistique* . Bruxelles: Pierre Mardaga, pp.421-489.

Paradis, M. 1989. Bilingual and polyglot aphasia. Dans F. Boller et J. Grafman (red.), *Handbook of Neuropsychology*, vol. 2. Amsterdam: Elsevier, pp. 117-140.

Paradis, M. 1993. Linguistic, psycholinguistic and neurolinguistic aspects of "interference" in bilingualspeakers: The Threshold Hypothesis. *International Journal of Psycholinguistics*, 9, 133-145.

Perecman, E. 1984. Spontaneous translation and language mixing in a polyglot aphasic. *Brain and Language*, 23, 43-63.

Rondal, J.A. 1994. Exceptional language development in mental retardation: Natural experiments in language modularity. *Cahiers de Psychologie Cognitive*, 13, 427-467.

Rondal, J. & A Comblain. 1995. Language in adults with Down Syndrome. Dans S. Buckley (réd.), *Language and cognitive development in Down's syndrome*. London: Chapman and Hall.

Weinreich, U. 1968. Unilinguisme et multilinguisme. Dans *Le Langage*, ouvrage sous la direction d'André Martinet (Encyclopédie de La Pléiade, No. 25). Paris: Gallimard. pp. 647-683.

Wray, A. 1992. *The focussing hypothesis*. Amsterdam: John Benjamins.

# Section 2

## *Sémantique lexicale et pragmatique*

Julie Auger et Yvan Rose (réd.). (1997). *Explorations du lexique*. Québec: CIRAL, 31-45.

# L'effet existentiel

*Denis Bouchard*
Université du Québec à Montréal

## 1. Des pronoms explétifs?

Cet article traite des pronoms "explétifs" *y* et *there* que l'on retrouve dans les constructions existentielles comme en (1):

(1) a Il y a une maison.
b There is a house.

Cela peut sembler curieux dans une section consacrée à la sémantique lexicale et à la pragmatique, puisqu'un explétif, par définition, n'a pas de sens. Mon but est en fait de soulever un point de méthodologie qui concerne directement ce thème, à savoir, la base sur laquelle on décide que telle paraphrase ou représentation correspond au sens d'une phrase. En général, le sens est supposé connu, presque évident: contrairement à une structure syntaxique par exemple, qu'on suppose ne se révéler à nous que par des tests multiples et sans cesse renouvelés, on pose le sens comme un donné et on tente de le relier de façon post hoc aux éléments et à la structure de l'élément visé.

Dans le cas de phrases comme celles en (1), en parlant de constructions *existentielles*, on pose justement que le sens de ces constructions est évident, typiquement, que ce sens est représenté par des primitifs empruntés à la logique (bien qu'ils ne soient compris que par analogie aux langues naturelles, comme l'a souligné Milner (1989)). La tâche consiste alors à trouver ce qui correspond à l'opérateur existentiel, à la variable qu'il lie, etc.

Je vais plutôt adopter une méthodologie qui tente de comprendre le sens de ces constructions à partir de ce qui est établi plus fermement à propos de leur syntaxe et du sens minimal de certains des mots en présence. Mais de toute façon, il faut noter que la méthodologie partant de la notion de construction existentielle n'est pas cohérente avec certains points majeurs de la théorie linguistique.

Premièrement, dans une approche minimaliste comme celle de Chomsky (1995), où il n'y a que deux niveaux de représentation, A-P (articulatoire-perceptuel) et C-I (conceptuel-intentionnel), parce qu'ils sont les seuls conceptuellement nécessaires, étant des interfaces, la présence même d'éléments explétifs dans la Grammaire est plutôt surprenante. Pourquoi une langue évoluerait-elle vers un état où elle en arrive à employer un explétif, s'il n'a aucune

pertinence à un des niveaux d'interface? En fait, si un vrai explétif existe, le linguiste théoricien doit trouver un moyen de le faire disparaître au niveau où se fait l'interprétation.[1] Deuxième incohérence, si on attribue un sens à la construction elle-même, cette conception holistique n'est pas cohérente avec la plupart des approches, la notion de construction n'ayant plus de statut théorique.

L'apport d'un explétif, vide de sens par définition, dans une phrase avec un sens existentiel demeure donc très énigmatique. Et pourtant, un grand nombre de langues construisent des phrases à "saveur existentielle" à partir de mots apparentés à *y* et à *there*. Je propose que ce qu'il y a de particulier à ces mots et qui en fait de bons candidats pour exprimer de l'existence, ce n'est pas d'être des explétifs, propriété qui leur est attribuée de façon post hoc, mais d'être des locatifs.[2] Qui plus est, ils se retrouvent dans une construction où on leur prédique les propriétés d'un SN.

## 2. Les éléments en présence

Regardons ces constructions, sans préjuger de ce qui devrait s'y trouver. Quels sont les éléments en présence, et comment sont-ils organisés entre eux? Commençons par l'anglais. Si on regarde la

[1] Ainsi, Chomsky (1995: 287) propose que la vérification raye le trait catégoriel d'un explétif, ce qui le rend invisible en forme logique, sans toutefois l'effacer parce qu'un terme de la numération ne peut disparaître complètement.

[2] Y et THERE sont dits explétifs parce qu'ils apparaissent dans une position où d'autres éléments sans théta peuvent apparaître, comme le sujet de *seem*, etc. Mais cet argument est circulaire, puisque la principale raison habituellement donnée pour la non-thématicité de cette position, c'est la présence possible d'explétifs. De plus, des éléments à théta peuvent apparaître dans cette position, comme dans *Jean semble avoir compris*. Mais on dira alors qu'ils n'ont pas de rôle théta parce que les restrictions de sélection ne sont que celles du verbe enchâssé. En fait, cela ne montre pas qu'il n'y a pas de rôle d'assigné, mais seulement que *sembler* est très peu contraignant, que tout peut sembler.
Un autre argument pour dire que Y et THERE sont explétifs repose sur le fait que leur référence dans ces emplois n'est pas claire. Mais il y a bien d'autres emplois où c'est le cas, mais où on ne dit pas qu'ils sont sans rôle thématique, comme dans les exemples suivants:

| | | |
|---|---|---|
| (i) | a | Il y va de votre vie. |
| | b | J'y compte bien. |
| (ii) | a | We're getting there (additionnant, etc) |
| | b | There I agree with you (argument) |
| | c | So there it is and nothing can be done about it. (situation) |
| | d | There you are (c'est ce que tu voulais, confirmation) |

surface de phrases comme (1b), on a *there* en position de sujet, suivi de la copule *be* et ensuite d'un SN dans la position où on trouve habituellement un SN prédicatif. La phrase s'apparente donc à des cas simples de prédication comme en (2) et (3).

(2) That is a house.
(3) a There is a Santa Claus in the play.
b John is a Santa Claus in the play.

L'observation que *there+be+*SN sont organisés comme dans une prédication simple, est loin d'être nouvelle. Jenkins (1972) proposait une structure de base pour les existentielles avec exactement cette propriété. Plus récemment, l'idée fait son chemin chez Safir (1982, 1989), Higginbotham (1989), Reuland et Meulen (1989:15), Keenan (1989: 300). Guéron (1986) étend l'idée au français: le locatif *y* et le SN *un homme* forment une petite proposition comme en (4b), où *y* occupe la position du sujet structural dans la SC et *un homme* occupe la position du prédicat.

(4) a Il y a un homme
b e avoir [SC [NP1 y] [NP2 un homme ]

Chez Jenkins, une existentielle est une prédication comme une autre. Pour Keenan et pour Reuland et Meulen, la prédication est spéciale parce que le *there* dénote un ensemble vide. Mais la plupart ne suivent pas ce que suggère si fortement la théorie syntaxique (probablement à cause d'une conception holistique de la construction). Ainsi, Safir, Higginbotham, Guéron reconnaissent la présence d'une relation prédicative, mais disent qu'il ne s'agit que d'une prédication "formelle" (conception qui semble contradictoire).

Il y a donc un certain consensus sur ce qu'on a dans ces construction: un SN est prédiqué d'un locatif (*y* , *there*). Mais s'agit-il vraiment d'un locatif? Ou est-ce un hasard qu'un homonyme d'un locatif apparaisse si souvent dans l'expression de l'existence? Une analyse basée sur un explétif voudrait qu'il s'agisse d'homonymie. Mais c'est là la pire des solution, qu'il ne faudrait adopter qu'en tout dernier recours, puisque la présence d'un locatif y est présentée comme un accident, qui plus est, un accident qui se reproduit très fréquemment d'une langue à l'autre.

De plus, la plupart des phrases avec *there* et *y* qui reçoivent une lecture existentielle, reçoivent aussi d'autres interprétations pour la même séquence de mots (mots qui semblent dans le même rapport syntaxique entre eux):

(5) a Il y a deux gardes.
a' Existentielle
a'' Deux gardes s'y trouvent. (déictique)
a''' Jean y maintient deux gardes. (déictique; *il* référentiel)

b There is a house.
b' existentielle
b'' Voilà une maison (une vraie) (présentationnelle)
b''' Là se trouve une maison (déictique)

L'effet présentationnel est plus net si le SN est défini et spécifique:

(6) Qui peut faire ce travail?
a Il y a Jean.
b There is John.

Comme le fait remarquer Chomsky (1995: 384, note 44), "the definiteness effect distinguishes interpretations, not well-formedness strictly speaking". Il en conclut que l'effet de défini est obscur. En fait, je ne crois pas qu'il y a d'effet de défini quand une phrase est existentielle. C'est plutôt l'inverse qui se passe: si un SN indéfini est prédiqué d'un locatif avec une référence dans un domaine mental, alors on en infère de l'existence. Reste à voir comment et pourquoi.

### 3. Le sens des existentielles

Dans une approche holistique aux existentielles, la première tâche consiste à déterminer ce que ça veut dire qu'une proposition fait une assertion d'existence. Une fois la réponse trouvée, il faut tenter de la relier aux propriétés morphosyntaxiques des constructions qui expriment ce concept. Par exemple, une réponse maintenant classique consiste à dire qu'il y a deux principaux sens (disparates, dit Milsark (1974, chapitre 6), s'appuyant sur Russell), à une assertion d'existence:

(7) Deux principaux sens de "assertion d'existence":
Plusieurs grands auteurs existent
Many great authors exist.
a EXISTER I: "a une réalité" (prédiqué d'un individu)
(ils ne sont pas le fruit de mon imagination)

b EXISTER II: "ensemble qui est instantié, qui a au moins un membre" (prédiqué d'un ensemble)
There are many great authors
Il y a plusieurs grands auteurs
(l'univers contient plusieurs grands auteurs, l'ensemble des grands auteurs contient plusieurs membres)

D'après Milsark, alors que le verbe *to exist* exprime le sens EXISTER I, les existentielles (*there* + BE) n'ont que le sens EXISTER II:

(8) There are many great authors.
Il y a plusieurs grands auteurs.

Toutefois, il faut remarquer que les impersonnelles avec *exister* n'ont également que le sens de EXISTER II:

(9) a There exist many great authors.
b Il existe plusieurs grands auteurs.

Le fait d'avoir une construction impersonnelle semble donc important, car même le verbe *exister* permet alors le sens de EXISTER II. En fait, les deux sens de "assertion d'existence" ne sont pas vraiment si disparates. Il y a deux facteurs dont il faut tenir compte. Premièrement, la nature +/- impersonnelle de la phrase détermine à propos de quoi est la phrase. On prédique toujours quelque chose à propos du sujet. Donc en (7a), on prédique quelque chose à propos du référent d'un SN, une classe préétablie de grands auteurs. Dans les termes de Voorst (1988), c'est ce SN sujet qui actualise l'événement, qui indique selon quelle perspective on saisit l'événement. Dans (7b), le SN *plusieurs grands auteurs* a perdu son rôle d'actualiseur: cette classe n'est pas préétablie, mais introduite dans le domaine conversationnel, d'où un effet présentationnel.

Le deuxième facteur concerne la référence aux mondes possibles implicite à toute phrase. Avec le verbe *exister*, cette référence semble restreinte à W1, "le monde réel"; ce serait donc une propriété lexicale de ce verbe. Dans les existentielles avec *il y a, there is*, la référence de l'ensemble "univers" peut être n'importe quel monde possible Wn. À cause de cette différence, nier avec *exister* n'entraîne pas de contradiction avec une construction existentielle: on nie que la chose soit dans le monde réel, mais pas dans tous les mondes possibles. Ainsi, supposons que l'oeuvre d'Homère et certains autres grands auteurs est le fruit du travail de plusieurs auteurs en fait, qu'il y a des

auteurs fictifs. On peut alors parfaitement dire (10) sans se contredire:

(10) Il y a de grands auteurs qui n'ont jamais existé.

En niant avec le verbe *exister*, on prédique une non-appartenance de grands auteurs au monde réel; ceci ne nous empêche pas de prédiquer leur appartenance à des mondes possibles (fictifs, erronés) à l'aide de l'expression *il y a SN*.

Les constructions existentielles et le verbe *exister* n'expriment pas deux sens disparates de EXISTER. Si on fait abstraction de l'effet présentationnel dû à la présence d'un sujet impersonnel, dans les deux cas, il y a un rapport avec un ensemble "univers"; c'est simplement la référence de celui-ci qui varie.

(11) a EXISTER I (verbe *exister*): "a une réalité"; référence de l'ensemble "univers"= W1, le monde réel.
b EXISTER II: référence de l'ensemble "univers"= Wn, un des mondes possibles, selon le contexte.

Cette référence à Wn dans les existentielles nous donne une clé pour comprendre la construction. Supposons que le locatif dans cette construction réfère à un monde possible. Adoptons aussi l'hypothèse minimale que nous sommes en présence d'une simple structure de prédication. La tâche ne consiste plus à apparier un opérateur existentiel quelconque avec une structure syntaxique. Il s'agit plutôt de comprendre les différentes lectures possibles pour ces structures prédicatives, et en particulier ce qui donne la lecture existentielle. Plus précisément, il faut comprendre pourquoi le fait de prédiquer certains SN de *there* ou *y* donne une lecture existentielle. Pour ce faire, il faut se rappeler ce qui se produit dans une prédication simple, d'une part, et comprendre la nature des pronoms *there* et *y*, d'autre part, afin de comprendre ce qui leur arrive quand ils sont le sujet d'une prédication. Soit la prédication simple en (12):

(12) Jean est un bon médecin.

Dans une analyse typique, on dira qu'un SN comme *Jean* dénote un ensemble de propriétés; la phrase en (12) est vraie si la propriétée dénotée par *bon médecin* est membre de l'ensemble des propriétés dénotées par *Jean*. Si une existentielle comme *There is a house* est une simple prédication, alors la phrase est vraie si la propriété dénotée par le SN *a house* est membre de l'ensemble des propriétés

dénotées par *there*. Pour juger, il faut déterminer ce que sont les propriétés de *a house* et de *there*.

Pour comprendre la nature du SN, je vais m'appuyer sur l'analyse de la distinction défini/indéfini que fait Heim (1982). Un défini comme *the house* se réfère à une entité dont l'existence dans le monde du discours fait partie des présupposés, tandis qu'un indéfini comme *a house* crée un nouveau référent. Quant à *there*, c'est l'anti-centre déictique, c'est à dire le complément du centre déictique, du MOI-ICI-MAINTENANT. En emploi spatial, le centre déictique est déterminé en prenant le corps du locuteur comme point de référence. Donc *there* réfère à un point dans l'espace identifié comme étant hors de l'espace immédiat du locuteur. L'anti-centre déictique *there* peut aussi avoir des emplois autres que l'emploi spatial (cf. la note 2), auquel cas ce n'est pas le corps du locuteur qui sert de point de référence, mais plutôt une propriété du locuteur pertinente à la dimension sémantique en jeu.[3]

Par exemple, dans les existentielles, au lieu d'un espace physique, c'est un espace mental qui est en jeu (voir Fauconnier (1984) à propos de cette notion) puisque *there* réfère à un monde possible Wn, une notion mentale. En emploi mental, on peut supposer que le centre déictique, c'est le locuteur en tant que Sujet de Conscience. Comme *there* est l'anti-centre déictique, il réfère à un point dans l'espace mental du locuteur, dans son Contenu de Conscience, mais autre que le locuteur, à tout autre point que le Sujet de Conscience lui-même. Quant au *y* du français, c'est un locatif très général, non déictique; en emploi mental, *y* réfère à l'espace mental en général, donc à tout le Contenu de Conscience.

Si on met ensemble les informations fournies par la nature du locatif et du SN, ainsi que par la structure prédicative, on obtient directement une explication de l'effet existentiel. Ainsi, dans une

---

[3] Que *there* puisse ainsi varier quant à sa façon de référer n'est pas plus extraordinaire que ce qui se passe avec d'autres expressions référentielles (cf. les travaux de Cantrall (1974), Coppieters (1982, 1990), Kuno (1987), Ruwet (1990), et bien d'autres). Par exemple, Coppieters observe qu'un SN peut être compris comme une entité-sujet, auquel cas le pronom de renvoi est *il*, ou comme une entité-concept, le pronom étant alors *ce*:

(i) Jean se rendait enfin compte qu'il avait tout essayé et tout raté. Il/C'était un homme pauvre désormais. Pourquoi avait-il eu tant de malchance?

Avec le pronom *il*, la question "Pourquoi avait-il eu tant de malchance?" exprime la réaction de l'entité-sujet *Jean*, tandis qu'avec *ce*, c'est une question rhétorique qui exprime l'opinion du locuteur.

phrase, le sujet étant l'actualiseur de l'événement (Voorst 1988), cet argument est donc typiquement du "vieux matériel", du connu, et ce qui en est prédiqué est du matériel nouveau.[4] Donc dans une phrase comme (13a), on a mis *there* référant au Contenu de Conscience, en position sujet, de "matériel connu", et un SN indéfini, nouveau référent, dans une position de "nouveau matériel".

(13) a There is a man outside (pas à propos de *a man*).
b Il y a un homme dehors (pas à propos de *un homme*).

Il en résulte un effet de présentation du SN, comme dans toute phrase prédicative, où la propriété du SN prédicatif est introduite dans le domaine du discours:

(14) Jean est un bon médecin.

Toutes les phrases avec un *there* initial ont un effet présentationnel à cause de cette particularité de prédiquer un SN d'un sujet *there* (la présence du *il* impersonnel a le même effet en français).

En prédiquant un SN de *y*, *there*, j'affirme que ce SN est une propriété de l'ensemble dénoté par *y* , *there*, donc de "points" dans mon Contenu de Conscience, comme par exemple un monde possible Wn.[5] Quand le SN est indéfini, comme dans les constructions dites existentielles, j'introduis donc un nouveau référent comme propriété d'un monde possible Wn qui a une identité dans mon Contenu de Conscience. Bref, à un point de l'espace mental du locuteur identifié comme étant Wn, celui-ci attribue la propriété 'un homme dans le jardin' ou 'a man'. C'est cette forme d'existence que j'affirme, une appartenance à mon Contenu de Conscience: c'est là l'effet existentiel. Il n'est nul besoin de chercher

---

[4] C'est pourquoi on n'a jamais de sujet [-déf, -spéc] et que même les indéfinis spécifiques sont bizarres en position sujet.
(i) #Un homme est dans le jardin

[5] Comme l'a fait remarquer Milsark (1974), dans la plupart des cas, la coda (la propriété prédiquée) n'est pas un SN, mais une structure plus complexe (on dirait aujourd'hui une phrase réduite "small clause"). Ainsi, une phrase réduite ne peut être sujet d'une phrase, mais peut être une coda:
(i) a There is a mosquito that can give you jungle rot.
b *A mosquito that can give you jungle rot is.
c There were several people arrested for public sanity.
d *Several people arrested for public sanity were.
Malgré ce fait, je vais continuer de référer au 'SN prédicatif', pour simplifier.

dans ces phrases un opérateur existentiel, qui de toute façon n'est compris qu'analogiquement.

**4. Effet d'indéfini, effet de liste, effet existentiel**

Dans notre analyse, il n'y a pas de construction existentielle à proprement parler, donc pas de construction qui exige la présence d'un SN indéfini. Il y a plutôt un effet existentiel qui survient quand les propriétés d'un SN d'un certain type sont prédiquées d'un locatif. Si le SN est défini, la phrase n'est pas agrammaticale, comme nous l'avons vu en (6) (repris ici), mais reçoit une autre lecture:

(15) Qui peut faire ce travail?
a Il y a Jean.
b There is John.

Le SN est alors présenté comme faisant partie d'une liste. Cet effet découle directement de notre analyse. En (15b), *there* réfère à un point de l'espace mental du locuteur et on attribue à ce point la propriété du défini *John*, c'est-à-dire la propriété d'une entité dont l'existence dans le monde du discours fait déjà partie des présupposés. Il y a donc contradiction, ou du moins forte redondance, si, comme dans le cas de la lecture existentielle, cet espace mental correspond à mon Contenu de Conscience conversationnellement pertinent. En effet, le défini fait déjà partie de ce contenu, des présupposés (cf. la discussion de Heim (1982) ci-dessus), mais se retrouve dans une position qui a pour effet de le présenter comme du nouveau matériel. Par contre, si *there* réfère plutôt à une liste pertinente pour le discours, alors il n'y a pas de contradiction. Le défini est présupposé faisant partie du domaine du discours, mais pas nécessairement de cette liste: le défini est du nouveau matériel par rapport à cette liste. Bref, le défini ici est connu, mais il n'est pas donné (Allan 1971: 13). L'effet de liste provient donc d'un choix différent pour la référence du locatif, une option obligatoire pour un défini afin d'éviter une contradiction, mais facultative pour un indéfini.

Selon cette analyse, ce n'est pas exactement le trait [+/- défini] du SN qui est en jeu dans l'effet de liste, mais plutôt le fait que le SN correspond à un référent dont les interlocuteurs sont conscients. Le lien avec le trait [+/- défini] provient du fait qu'un SN défini réfère généralement à une entité spécifique dont les interlocuteurs sont conscients. Toutefois, un défini non-spécifique ne devrait pas induire de contradiction avec une lecture existentielle, selon notre analyse.

C'est effectivement ce qu'on observe dans des exemples comme en (16), tirés de Lakoff (1987):

(16) There was the usual argument in class today.
There's the strangest smell coming from the refrigerator.
There's this woman next door who plays the stereo too loud.

Ici, *the* et *this* ne sont pas utilisés pour identifier des référents spécifiques dont mon interlocuteur serait déjà conscient. Il n'y a donc pas de contradiction avec une lecture existentielle puisqu'on n'attribue pas au Contenu de Conscience conversationnellement pertinent du matériel qui serait déjà présupposé.

En résumé, une construction dans laquelle un SN est prédiqué de locatifs comme *y* et *there* peut induire un effet existentiel ou un effet de liste selon la référence du locatif et du SN.

## 5. Le choix du pronom locatif

Dans cette analyse, il n'est pas accidentel qu'un locatif se retrouve si souvent dans les constructions qui servent à exprimer le type d'assertion existentielle discuté dans la section 3. Il doit y avoir un élément qui puisse référer à un lieu mental, d'où la présence d'un locatif. Il reste à déterminer pourquoi ce sont les locatifs *y* et *there* qui sont utilisés en français et en anglais pour faire cette assertion d'existence, et non pas d'autres locatifs.

Ce type d'assertion d'existence maximalement générale se fait en attribuant les propriétés d'un SN à la plus vaste partie possible du Contenu de Conscience. Le locatif qui réfère au Contenu de Conscience dans cet emploi doit donc être le moins restrictif de la langue, le plus englobant, afin de couvrir le plus possible de l'espace mental pour être maximalement général, procurant ainsi une inférence "existentielle". En anglais, le système des pronoms locatifs ne comprend que deux éléments:

(17) Système des pronoms locatifs non-interrogatifs en anglais:
Centre déictique, proximité ("l'ici"): here
Anti-centre déictique, éloignement ("l'ailleurs"): there

Comme "l'ailleurs" est plus vaste, général, que "l'ici", c'est donc le pronom *there* qui sert à exprimer une assertion d'existence maximalement générale en anglais. En français, le système est légèrement plus complexe:

(18) Système des pronoms locatifs non-interrogatifs en français:

| | |
|---|---|
| Centre déictique, proximité: | ici |
| Anti-centre déictique, éloignement: | là-bas |
| Déictique neutre:[6] | là |
| Locatif pur, sans déicticité: | y |

Le locatif le moins marqué en français étant *y*, c'est ce pronom qui sert à exprimer une assertion d'existence maximalement générale.

Certaines différences entre la contruction existentielle du français et celle de l'anglais découlent de ces choix de pronom. Premièrement, le locatif des existentielles est un objet en français (19a), tandis qu'il doit être un sujet en anglais (19b):

(19) a Il y a un problème.
a' *Y est un problème.
b There is a man.
b' A man is there (pas existentielle)

L'impossibilité d'avoir une lecture existentielle en (19b') découle du fait que le locatif doit être un argument du SN prédicatif, et non l'inverse, parce que l'assertion d'existence est inférée de l'attribution des propriétés du SN au Contenu de Conscience auquel réfère le locatif. L'anglais peut utiliser la structure prédicative simple LOC *be* SN comme en (19b), parce que le locatif le plus général *there* est un sujet potentiel. Mais le *y* du français est un clitique objet, et ne peut donc apparaître en position de sujet comme en (19a'). Il faut donc employer une construction prédicative de type petite proposition comme en (19a), où *y* occupe la position du sujet structural dans la SC et le SN occupe la position du prédicat (cf. (4)).

On comprend alors pourquoi on trouve une deuxième différence entre les deux langues: il y a un *il* impersonnel en français en plus du locatif *y*, mais seulement le locatif *there* en anglais. Ceci découle du besoin d'un sujet dans la phrase française, que le pronom *y* ne peut satisfaire parce qu'il est un clitique objet.[7]

Le fait que les différents emplois de *there* ne se traduisent pas par le même élément en français (et dans beaucoup de langues) est souvent présenté dans les études comparatives comme un argument

---

6 *Là* est déictique, mais neutre quant à la proximité ou l'éloignement:

(i) a Mets-toi là
b Viens là

7 Quant au *il* "explétif", il faudra une étude approfondie pour en déterminer l'apport sémantique, qui, je crois, n'est pas nul.

contre une analyse où il n'y a qu'un seul élément *there* dans tous les emplois (existentiel, déictique, présentationnel, etc.).

(20) *There* is a Santa Claus in the garden.
Il y a un Père Noël dans le jardin.
(21) I put them *there*.
Je les ai mis *là*.
(22) *There*'s a fine argument.
*Voilà* un bon argument.

Mais nous avons vu qu'il n'est pas nécessaire de postuler plusieurs lexèmes *there* en anglais. Ces différences découlent simplement du fait qu'il n'y a pas le même paradigme de pronoms locatifs en français, qu'il n'y a pas de pronom français qui correspond à *there*. Pour référer aux différentes situations des exemples (20) à (22), le français emploiera les pronoms appropriés dans son paradigme.

Reuland et Meulen (1989: 6) apportent une autre objection de nature comparative à une analyse des existentielles qui serait basée sur la présence d'un lexème particulier dans la construction. Comme il y a des langues avec des phrases existentielles qui n'ont pas d'explétif sujet en surface, la seule approche plausible pour réconcilier ces diverses analyses, selon eux, est de faire dériver les propriétés attribuées à *there* d'autres propriétés de la construction. Ainsi, en chinois, les existentielles sont construites avec AVOIR sans qu'il n'y ait de pronom en surface:

(23) You gui Huang (1989)
avoir fantômes
Il y a des fantômes (ici)

On pourrait croire que leur argument pour une approche holistique s'adresse tout aussi bien à mon analyse, puisque j'attribue également la lecture existentielle à la présence d'un lexème particulier, non pas un explétif, mais un locatif auquel on prédique les propriétés d'un SN. Comment analyser alors une langue comme le chinois, où il n'y a apparemment ni explétif, ni locatif dans les existentielles?

D'abord, il faut préciser que je ne dis pas qu'employer un locatif général auquel on prédique les propriétés d'un SN est la seule façon d'exprimer l'existence (le verbe *exister* du français le fait tout autant, avec des nuances). Mais le fait est qu'on retrouve en chinois un verbe au sens très général, proche du verbe *avoir*, qui peut

apparemment exprimer l'existence sans l'aide d'un locatif. Ce contraste entre les deux langues exige une réponse plus détaillée.

Certaines données additionnelles du chinois sont très éclairantes. Comme on le voit en (24) et (25), un syntagme locatif peut apparaître en position sujet du verbe *you* 'avoir'.

(24) Zhuo-shang you yiben shu
table-dessus avoir un livre
Sur la table il y a un livre

(25) Zhuo-shang you yiben shu hen youqu
table-dessus avoir un livre très intéressant
Sur la table il y a un livre très intéressant

De plus, on sait que le chinois est une langue où le sujet n'a pas à apparaître en surface.[8] La présence de constructions existentielles comme (23) en chinois n'a donc rien de surprenant dans le cadre de l'analyse proposée ici. Selon cette analyse, c'est le pronom le moins restrictif de la langue qui est employé dans les constructions existentielles basées sur la prédication. En anglais, *there* est le pronom le moins restrictif, "l'ailleurs" de *here*; en français, c'est *y*, le locatif non déictique. Le chinois dispose d'un pronom encore plus général, un pronom zéro, sans aucune restriction. Il y a toutes les raisons de croire que ce pronom en position sujet de *you* peut référer comme un locatif, puisque *you* admet des syntagmes locatifs comme sujet. Ces trois langues semblent donc pouvoir exprimer l'existence à l'aide d'un locatif général et d'un SN prédicatif. On évite ainsi les problèmes soulevés par une approche holistique qui attribuerait des propriétés existentielles à la construction.

## 6. Conclusion

La syntaxe élémentaire des phrases existentielles indique que la construction existentielle est formée d'un élément locatif en position d'argument externe d'un SN prédicatif. L'effet existentiel provient d'une combinaison de propriétés particulières de ces éléments, plus précisément, lorsque l'élément locatif fait référence au Contenu de Conscience et que le SN prédicatif n'est pas à la fois défini et spécifique. Le sens d'une existentielle, c'est d'attribuer les propriétés

8 Il n'est pas directement pertinent pour mon analyse de déterminer si cette absence de sujet est du même type que ce que l'on retrouve dans certaines langues romanes, ou est une topicalisation nulle comme le propose Huang (1982).

de ce SN au Contenu de Conscience, d'où l'on tire une inférence d'existence.

Selon cette analyse, ce n'est pas un accident que les locatifs servent si souvent à exprimer l'existence dans les langues. Le choix du locatif n'est pas accidentel non plus: il faut que ce soit le plus englobant possible. La relation de prédication inhabituelle, où c'est le locatif qui est l'argument du SN plutôt que l'inverse, s'explique aussi.

On n'a pas à postuler de pronom explétif, ni à attribuer un sens de façon holistique à la construction, évitant ainsi des problèmes de cohérence avec la théorie générale. Point n'est besoin non plus d'assigner de multiples structures aux phrases en (5) selon les différentes lectures qu'elles peuvent recevoir. Il n'y a que la référence du locatif et les présupposés rattachés au SN qui varient selon le contexte. Et cette variation n'a rien d'extraordinaire: on la retrouve pour toutes les expressions référentielles.

En somme, si on se donne comme tâche non pas d'apparier un opérateur existentiel de façon holistique avec une structure syntaxique, mais plutôt de comprendre les différentes lectures possibles pour ces structures prédicatives, et en particulier la lecture existentielle, on se rend compte encore une fois de la remarquable économie de moyens du langage.

**Bibliographie**

Allan, Keith. 1971. "A Note on the Source of There in Existential Sentences". *Foundations of Language* 7: 1-18.

Cantrall, William R. 1974. *Viewpoint, Reflexives, and the Nature of Noun Phrases*. The Hague: Mouton.

Chomsky, Noam. 1995. *The Minimalist Program* . Cambridge: MIT Press.

Coppieters, René. 1982. "Descriptions and attitudes: The problem of reference to individuals". *Studies in Language* 6: 1-22.

Coppieters, René. 1990. "Individuals, Thematic Roles, and Semantic Interpretation". ms., Pomona College.

Fauconnier, Gilles. 1984. *Espaces mentaux* . Paris: Editions de Minuit.

Guéron, Jacqueline. 1986. "Le verbe *avoir* ". ms., Université Paris VIII.

Heim, Irene. 1982. *The Semantics of Definite and Indefinite Noun Phrases* . Thèse de doctorat, University of Massachusetts, Amherst.

Higginbotham, James. 1989. "Indefiniteness and Predication". In E. Reuland & A. ter Meulen (réd.), p.43-70.

Huang, C.-T. J. 1982. *Logical Relations in Chinese and the Theory of Grammar*. Thèse de doctorat, Massachusetts Institute of Technology.

Huang, C.-T. J. 1989. "Existential Sentences in Chinese and (In) definiteness". In E. Reuland and A. ter Meulen (réd.), p. 226-253.

Jenkins, Lyle. 1972. *Modality in English Syntax* . Thèse de doctorat, Massachusetts Institute of Technology.

Keenan, Edward. 1989. "A Semantic Definition of "Indefinite NP"". In E. Reuland and A. ter Meulen (réd.), p. 286-317.

Kuno, Susumu. 1987. *Functional Syntax: Anaphora, Discourse and Empathy* . Chicago: The University of Chicago Press.

Lakoff, George. 1987. *Women, Fire, and Dangerous Things* . Chicago: The University of Chicago Press.

Milner, Jean-Claude. 1989. *Introduction à une science du langage* . Paris: Editions du Seuil.

Milsark, Gary. 1974. *Existential Sentences in English* . Indiana University Linguistics Club.

Reuland, Eric & Alice ter Meulen. 1989. *The Representation of (in)definiteness*. Cambridge: MIT Press.

Ruwet, Nicolas. 1990. "EN et Y: Deux clitiques pronominaux antilogophoriques". *Langages* 97: 51-81.

Safir, Kenneth. 1982. *Syntactic Chains and the Definiteness Effect*. Thèse de doctorat, Massachusetts Institute of Technology.

Safir, Kenneth. 1989. "What Explains the Definiteness Effect?" In E. Reuland and A. ter Meulen (réd.), p.71-97.

Voorst, Jan van. 1988. *Event Structure* . Amsterdam: Benjamins.

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 47-62.

# Fonctions configurationnelles[1]

Anna-Maria Di Sciullo, UQAM

## 1. Introduction

Prenant pour acquis que la connaissance lexicale est très limitée à l'état initial, nous proposons que la grammaire inclut une fonction qui permet d'obtenir des catégories dérivées à partir de catégories plus simples. Notre proposition permet d'expliquer pourquoi certains termes individuels, en particulier les affixes dérivationnels, ont à la fois des propriétés lexicales et des propriétés fonctionnelles. Ces termes ont des propriétés lexicales puisqu'ils dérivent des catégories lexicales, des noms (N), des verbes (V) et des adjectifs (A). Ils ont aussi des propriétés fonctionnelles telles que définies par exemple dans Abney (1987) et Szabolsci (1987). Comme les autres catégories fonctionnelles, telles que les quantifieurs (Q) et les déterminants (D), les affixes dérivationnels sont des items appartenant à des classes finies et bien qu'ils dénotent, ils ne décrivent pas. Pourquoi en est-il ainsi?

Nous adoptons une approche configurationnelle à la question de savoir pourquoi certains termes ont plus d'une valeur catégorielle associée et proposons que la grammaire inclut une fonction configurationnelle F($x$). Cette fonction dérive des catégories étendues, i.e. des structures de traits complexes, et des prédicats étendus, i.e. des structures prédicat-arguments complexes. Le rôle de cette fonction est de contribuer à l'économie de la grammaire.

Cet article est organisé comme suit. Nous discutons en premier lieu de la formation des catégories complexes en montrant que la catégorie de certains affixes est établie configurationnelle-

---

[1] Des versions préliminaires de ce travail ont été présentées à l'Université McGill en mai 1996 dans le cadre du Colloque sur le lexique (ACFAS) et à l'Université de Gérone en juillet 1996 dans le cadre du Colloque organisé par l'École Internationale de Linguistique de Gérone (GLOW). Je remercie les participants pour leurs questions et leurs interventions. Ce travail est subventionné par le Conseil de recherche en sciences humaines du Canada (411-92-0012) et par le Fonds pour la Formation de Chercheurs et l'aide à la Recherche (95-ER-0401).

ment. Nous proposons que F(*x*) dérive des catégories étendues, maintenant au minimum les traits catégoriels disponibles universellement à l'état initial de la grammaire. Nous discutons ensuite de la formation de structures prédicat-arguments complexes à partir de l'examen de formes verbales qui permettent ou excluent l'alternance ergative/transitive.

## 2. Catégories étendues

### 2.1 Affixes, sous-spécification et positions

Les traits catégoriels de certains affixes varient selon leur position dans la projection dont ils font partie. Ainsi, dans les verbes dénominaux et déadjectivaux de l'anglais, le même affixe en peut occuper plus d'une position.

(1) a. to lengthen the hem/to whiten the wall
b. to encode the message/to enlarge the road
c. to enliven the creature/to enlighten the issue

Prenant pour aquis la Règle de la Tête à Droite (cf. Williams 1981; Di Sciullo et Williams, 1987), l'affixe -en en position finale a des traits verbaux. Assumant le principe d'Identification des Adjoints (cf. Di Sciullo, 1996b), l'affixe en- en position initiale a des traits prépositionnels. Ainsi, en (2a), -en a des traits verbaux, qui ne peuvent être projetés à partir du nom ou de l'adjectif avec lequel il se compose. En (2b), en- n'a pas de traits verbaux, puisque dans plusieurs cas, aucun préfixe n'est requis dans la structure verbale, qui a nonobstant des traits verbaux, (3), (cf. Di Sciullo, 1996a). Donc, en (2c), le préfixe a donc des traits prépositionnels et le suffixe, des traits verbaux.

(2) a. lighthen b. enlarge c. enlighten
(3) to code the message/to list the ingredients/to lighten the color

Dans la perspective de réduire l'information lexicale à son minimum, et donc d'exclure des entrées multiples pour un même terme, l'affixe en par exemple, nous posons que la sous-spéci-

fication fait partie de tous les systèmes de traits manipulés par la grammaire, non seulement des systèmes de traits phonologiques et morphologiques (cf. Archangeli, 1984; Archangeli et Pulleyblank, 1994; Lumsden, 1992; Rooryck, 1994), mais aussi des systèmes de traits catégoriels et sémantiques. Les entrées lexicales incluent des traits sous-spécifiés, et la spécification lexicale est réduite au minimum. Soit les définitions suivantes.

(4) Lexique$_{def}$ = ensemble de spécifications {LF,PF} non dérivables
(5) Spécifications lexicales$_{def}$ = {ensemble de traits}
(6) Sous-spécification lexicale$_{def}$ = seules les valeurs marquées sont spécifiées

Prenant pour acquis le système de traits lexicaux de Chomsky (1970), (7), et notre hypothèse selon laquelle les traits catégoriels peuvent être sous-spécifiés, il est possible de rendre compte des faits notés plus haut en posant que l'affixe en n'a aucune valeur pour le trait [V], et une valeur négative pour le trait [N], (8). La valeur de [V] est dérivée dans les configurations par la Règle de la Tête à Droite et le Principe d'Identification des Adjoints.

(7) V: [-N, +V]; N: [+N, -V]: A: [+N, +V]; P: [-N -V]
(8) en: [-N, V]
(9) a. [light$_A$ -en$_V$]; b. [[en-$_P$ large$_A$ ]$_V$]; c. [[en-$_P$ large$_A$] -en$_V$]

Nous posons donc que la sous-spécification couvre aussi le système des traits catégoriels et que la catégorie de certains termes est dérivée. Ceci permet d'éviter la reduplication des entrées lexicales pour des termes singuliers, et donc de restreindre l'information marquée dans le lexique.

## 2.2 **Traits et domaines**

### 2.2.1. **P/ASP**

Notre approche à l'identification des traits des termes singuliers est motivée par le fait que l'interprétation de certains

préfixes du français est une fonction du domaine dont ils font partie. Ainsi, considérant le préfixe dé- du français, nous observons que ce préfixe est interprété comme une préposition directionnelle en (10a) et comme une préposition temporelle en (10b), supposant que les adverbes ne constituent pas une catégorie distincte des prépositions ou des adjectifs (cf. Emonds, 1985).

(10) a. débarquer b. défaire

Des différences d'ordre linéaire, de séparabilité et de structure d'argument appuient l'hypothèse que la spécification de traits des affixes est une fonction de leur position dans un domaine. Le préfixe directionnel dé- doit suivre un préfixe temporel, tel que le préfixe itératif re-, qui est à l'extérieur de la structure d'argument du verbe (Di Sciullo et Klipple, 1994; Koopman, 1995). De plus, le préfixe directionnel dé- est inséparable de la structure verbale, alors que le préfixe temporel dé- est séparable:

(11) redébarquer/*dérebarquer de l'avion
(12) débarquer/embarquer/*barquer de l'avion
(13) défaire/refaire/faire la valise

Le préfixe directionnel contribue à la structure d'argument de la projection verbale à laquelle il participe autorisant un argument locatif, alors que ce n'est pas le cas du préfixe temporel.

(14) débarquer DP de la voiture/embarquer DP en voiture
(15) défaire un puzzle/faire un puzzle

Le préfixe directionnel dé- est en distribution complémentaire avec d'autres préfixes directionnels, incluant a- et en-, ce qui n'est pas le cas du préfixe temporel re-.

(16) réengainer/*désengainer; redéterrer/*désenterrer; raccrocher/*déaccrocher

Ces faits nous conduisent à poser que le préfixe dé- peut être projeté soit à l'intérieur ou à l'extérieur de la structure d'argument lexicale du verbe. S'il est projeté à l'intérieur, (17a), il a les traits des prépositions directionnelles, s'il est projeté à l'extérieur, (17b) il a les traits des prépositions temporelles.

(17) a. [[P N] V ] b. [P [ V ]]

Cette différence configurationnelle est motivée également par le fait que le dé- interne contribue à 'l'aktionsart', c'est-à-dire la structure interne de l'événement dénoté par le verbe; alors que le dé- externe affecte la totalité de l'événement et non sa structure interne. Ainsi, en (18), le préfixe change une activité en un accomplissement, ce qui n'est pas le cas en (19).

(18) a. Il a placé le sofa *pendant cinq minutes/en cinq minutes.
b. Il a déplacé le sofa pendant cinq minutes/en cinq minutes.
(19) a. Elle a fait la chaise *pendant une heure/en une heure.
b. Elle a défait la chaise *pendant une heure/en une heure.

Ces faits sont consistants avec la généralisation que les préfixes externes tel que le préfixe itératif re-, n'affectent pas les traits aspectuels de la projection verbale à laquelle ils appartiennent, alors que les préfixes internes, tels que a- et en-, peuvent le faire (Di Sciullo, 1996b).

(20) a. Il a couru (pendant 5 minutes/*en cinq minutes).
b. Il est accouru (*pendant cinq minutes/en cinq minutes).
c. Il a recouru (pendant 5 minutes/*en cinq minutes).
(21) a. Il les a portés pendant une heure/#en une heure.
b. Il les a en/apportés #pendant une heure/en une heure.
c. Il les a reportés pendant une heure/#en une heure.

Si ASP est une projection indépendante (cf. Hendrick, 1991; Koopman, 1995; Travis, 1992) et si les traits aspectuels, tels que les traits terminatif [T] et sous-intervalle [S] font partie de la

spécification des catégories verbales et prépositionnelles (cf. Di Sciullo, 1996a), les préfixes ont des traits prépositionnels et aspectuels donnés selon leur domaine d'interprétation.

Ainsi, en (22a), le préfixe identifie un trait aspectuel du verbe auquel il est adjoint. En (22b), le préfixe identifie un trait aspectuel de l'événement entier dénoté par la projection verbale à laquelle il est adjoint.

(22) a. [P/ASP V] b. [P/ASP [V ]]

Le préfixe dé- peut donc être projeté à l'intérieur ou à l'extérieur de la structure d'argument verbale. S'il est projeté à l'intérieur, il est interprété comme une préposition directionnelle, s'il est projeté à l'extérieur, il est interprété comme une préposition temporelle.

2.4 **La fonction F(*x*)**

Prenant pour acquis que les domaines minimaux et étendus (cf. Chomsky, 1995) sont donnés à l'état initial de la grammaire, nous proposons l'existence de la fonction F(*x*) qui prend une catégorie dans un domaine minimal et donne une catégorie complexe dans un domaine étendu.

(23) $[_Z$ W $[_X$ X Y]]
Z est soit X ou différent de X

(24) F(*x*) = la valeur de *x* dans un domaine étendu

Des termes singuliers, tels que l'affixe dé-, sont sujets à cette fonction, puisque la valeur catégorielle de cet affixe dépend de son inclusion dans le domaine minimal ou étendu d'une tête verbale.

(25) dé-: [-V,-N] = P F[-V,-N] = $P_E$

Plus généralement, F(*x*) est nécessaire étant donné l'existence de termes dont la valeur catégorielle est une fonction de leur

position dans le domaine minimal ou étendu de la tête à laquelle ils sont associés. C'est le cas notamment de l'ensemble des affixes dérivationnels. Ainsi, le suffixe *-er* de l'anglais a à la fois des traits nominaux, dérivant des noms à partir de verbes, (26a), et des traits de quantificateur, en ce qu'il lie la variable argumentale externe de la projection verbale avec laquelle il se compose; ce suffixe peut se composer avec des verbes inergatifs mais pas avec des verbes ergatifs, (26b), (cf. Di Sciullo, 1996b).

(26) a. He writes operas./He is an opera writer.
b. He is a worker./*He is an arriver.

Le suffixe *-able* a des traits adjectivaux, puisqu'il dérive des adjectifs à partir de verbes; il a aussi des traits modaux puisqu'il ne peut se combiner avec des modaux, et qu'il dérive des prédicats de niveau individuel (individual level) à partir de prédicats à stade (stage-level). Comme c'est le cas des prédicats de niveau individuel, les adjectifs déverbaux en *-able* ne peuvent faire partie de constructions explétives et des expressions temporelles ou spatiales ne peuvent les modifier (cf. Di Sciullo, 1995).

(27) a. export/exportable/These goods are exportable.
b. They are able to export these goods.
(28) a. It is possible/necessary./*It must can be./*It is mayable
b. French is acquirable/*French is knowable.
(29) a. *There are chairs transformable./There are chairs available.
b. *There are linguists hospitable at the inn/this morning.
c. *Two kilos are weighable./This book is weighable.

Le suffixe verbal *-en* a des traits verbaux et des traits d'auxiliaire. Il permet de dériver des verbes à partir d'adjectifs, ainsi que des formes participiales.

(30) a. He boldens the title./*He bolds the title.
b. He has eaten./*He has eat./*He eaten.

## 2.5 Couverture de F(*x*)

Supposant que les deux traits binaires [N] et [V] définissent la classe des catégories lexicales et que les traits lexicaux sont sous-spécifiés, chaque catégorie lexicale soumise à F(*x*) devient une catégorie étendue. F(*x*) donne la valeur étendue de cette catégorie. De sorte que nous obtenons ce qui suit.

(31) a. [-V, -N ] P F(P) = $P_E$
b. [-V, +N ] N F(N) = $N_E$
c. [+V, +N ] A F(A) = $A_E$
d. [+V, -N ] V F(V) = $V_E$

Ainsi, une préposition P soumise à F(*x*) devient une préposition étendue, $P_E$, soit une marque d'aspect (ASP). La fonction F(*x*) permet de dériver les valeurs étendues pour les noms, soit Q, pour les adjectifs, soit MOD et pour les verbes, soit AUX.

Supposant que les deux traits binaires [D] et [T] définissent les catégories fonctionnelles, (cf. Chomsky, 1995), F(*x*) donne la valeur étendue de ces catégories. Nous obtenons ce qui suit.

(32) a. [+D, -T ] D F(D) = $D_E$
b. [-D, +T] T F(T) = $T_E$
c. [-D, -T] ASP F(ASP) = $ASP_E$
d. [+D, +T ] C F(C) = $C_E$

Ainsi, un pronom (D) non branchant soumis à F(*x*) est un déterminant étendu ($D_E$), soit un pronom clitique, comme c'est le cas de le, la, les en français. La fonction F(*x*) permet de plus de dériver des valeurs étendues pour le temps ($T_E$), pour l'aspect ($ASP_E$) et pour le complémenteur ($C_E$). Les traits des catégories étendues ne sont pas spécifiés lexicalement mais dérivés par la fonction F(*x*).

Le système proposé a les avantages théoriques suivants. Il restreint l'ensemble des catégories lexicales et fonctionnels aux traits [V], [N] et [D], [T], et il permet une définition configurationnelle des catégories dérivées qui sont des propriétés en partie

lexicales et en partie fonctionnelles. De plus, il est basé sur des propriétés indépendantes de la grammaire, soit la sous-spécification et les projections minimales et étendues.

2.6 **Résumé**

F(*x*) fournit ainsi une solution configurationnelle à la question de savoir pourquoi certains termes sont associés à plus d'une catégorie et pourquoi certaines catégories font partie d'ensembles restreints.

3. **Prédicats étendus**

La notion de fonction configurationnelle permet aussi de fournir une solution configurationnelle aux phénomènes d'alternances transitive/intransitive, possibles pour certains prédicats verbaux.

3.1. **Les affixes verbaux**

Les verbes déadjectivaux de l'anglais incluant le suffixe -en sont intransitifs/inchoatifs et dénotent des changements d'état. Ils peuvent généralement être transitivisés.

(34) The apple reddens./The sun reddens the apples.
(35) The knife sharpens./The butcher sharpens the knife.

Similairement, en français, les verbes dérivés en -ir sont intransitifs/inchoatifs; l'adjectif qu'ils incluent exprime l'état résultant putatif de l'événement. Ces verbes peuvent aussi être transitivisés.

(36) La peau brunit./Le soleil brunit la peau.
(37) Le fer rougit./Le forgeron rougit le fer.

Ces affixes diffèrent des affixes -ize de l'anglais et -er du français qui font partie de verbes dénominaux et déadjectivaux qui sont transitifs/causatifs. La catégorie nominale qu'ils incluent exprime le lieu ou l'état résultant.

(38) He rationalized the production./*He rationalized.
(39) Il a affiné l'instrument./*Il a affiné.

Que certains affixes verbaux puissent projeter une structure d'argument intransitive aussi bien que transitive est un effet de la fonction F(*x*).

3.2. **F(*x*)**

Prenant pour acquis que la projection asymétrique en (40) est une structure d'argument élémentaire, (cf. Hale & Keyser, 1993), nous supposerons que la position argumentale W est sous-spécifiée. Cette position peut être soit argumentale, soit non-argumentale, contrairement à la position Y qui est argumentale. Nous proposons que la fonction configurationnelle F(*x*) permet de dériver une valeur argumentale [+A] pour la position W, la valeur par défaut étant [-A].

(40) $[_X \underline{W} [_X X Y ]]$

La fonction F(*x*) rend compte de la formation de structures prédicat-arguments complexes à partir de prédicats élémentaires. Les affixes verbaux -en et -ir sont sous-spécifiés pour [A] et sont sujets à cette fonction. Les autres affixes verbaux dans ces langues ne sont pas sujets à cette fonction, étant spécifiés lexicalement [+A] pour la position W.

(41) a. -ize : [+A] (anglais) c. -e(r) : [+A] (français)
b. -en : [A ] d. -i(r) : [A ]

L'ensemble des affixes verbaux dans ces langues est limité à deux, puisque les affixes diffèrent quant à la valeur [A] de l'unique position W.

### 3.3 **Couverture de F(*x*)**

Nous supposerons que les prédicats élémentaires sont monadiques, i.e. ils n'ont qu'un seul argument dans leur domaine. F(*x*) étend le domaine d'un prédicat de telle sorte qu'il inclut un autre argument. F(*x*) couvre la formation de prédicats complexes, que le prédicat soit un affixe, comme c'est le cas pour certains affixes verbaux, ou une forme libre, comme c'est le cas de la formation de prédicats transitifs à partir de verbes ergatifs, (42). Supposant maintenant que la position sous-spécifiée vis-à-vis de [A] est la position Y en (40), F(*x*) couvre également la formation de prédicats transitifs à partir de prédicats inergatifs, (43).

(42) a. The clock broke./John broke the clock.
b. Les prix ont augmenté./Ils ont augmenté les prix.
(43) a. He spoke./He spoke their language.
b. Il a parlé./Il a parlé leur langue.

#### 3.3.1 **L'alternance ergative/transitive**

L'alternance ergative/transitive est discutée notamment dans van Valin (1987) et Grimshaw (1990), qui fournissent un traitement aspectuel du phénomène, ainsi que dans Levin et Rappaport (1995), qui présentent une approche sémantique lexicale qui fait appel à des notions de contrôle interne et externe de l'événement dénoté par le verbe. L'alternance ergative/transitive est généralement vue comme un cas d'alternance activité/accomplissement.

(44) a. accomplissement [[activité] [changement d'état]] — transitif
b. [activité]] — inergatif

Considérant les faits du français, Houle (en préparation) montre que l'alternance transitive/ergative est possible avec certains verbes de chaque classe aspectuelle de verbes, i.e. les états, les activités, les accomplissements et les achèvements, tels que définis traditionnellement (cf. Vendler, 1967). De sorte qu'il est nécessaire de trouver une autre explication au phénomène de l'alternance

transitive/ergative, qui ne peut être réduit à une alternance entre deux classes aspectuelles de verbes.

| (45) Accomplissement **Transitif** | Accomplissement **Intransitif** |
|---|---|
| a. Il a changé son apparence. | a'. Son apparence a changé. |
| Accomplissement | État |
| b. Il dévie la route. | b'. La route dévie. |
| Achèvement | Achèvement |
| c. Il casse le vase. | c'. Le vase casse. |
| Achèvement | État |
| d. Il pend sa veste au crochet. | d'. Sa veste pend au crochet. |
| Activité | État |
| e. Il trempe sa plume. | e'. Sa plume trempe. |
| Activité | Activité |
| f. Il roule le baril. | f'. Le baril roule. |

Si l'alternance ergative/transitive n'est pas uniquement un phénomène aspectuel, il est possible de fournir un traitement naturel de ce phénomène en termes de la fonction configurationnelle F($x$) et de la sous-spécification des traits sémantiques.

(46) $[_{v}\ W\ [_{v}\ V_{[\ ]}\ Y]]$

Nous proposons que les verbes inaccusatifs qui sont sujets à la fonction F($x$) sont sous-spécifiés pour le trait inhérent [I], discuté dans Houle (en préparation). [I] fait partie des spécifications de traits conceptuels pour les catégories conceptuelles Inchoatif [B] et Causatif [D], que nous définissons comme suit.

(47) Inchoatif : [+B, -D, +I] Causatif : [-B, +D, -I]

Le trait [I] spécifie si le changement d'état résultant est inhérent à l'événement ou non: si [+I], le changement est inhérent et si [-I], le changement est l'effet d'un facteur externe.

(48) [+I] : changement d'état inhérent
[-I] : changement d'état non inhérent

La modification d'un nominal d'événement par des expressions incluant des réfléchis donne lieu à des interprétations différentes si le verbe de base est [+I] ou [-I]. Dans le premier cas, cela donne lieu à une interprétation redondante et dans le second cas cela donne lieu à une contradiction.

(49) a. #The falling of the snow happened by itself.
b. #The arrival/leaving of John happened by itself.

Les verbes inaccusatifs [+I], tels que le verbe to die 'mourir' en anglais et naître en français, ne sont pas sujets à F(*x*), ainsi que les verbes inergatifs [-I] tel que le verbe arrive 'arriver' en anglais et le verbe partir en français.

(50) He died./*He died Peter./ Il est né./*Il est né Paul.
He arrived./*He arrived John./Il est parti./*Il est parti Paul.

Seuls les verbes inaccusatifs qui sont sous-spécifiés pour le trait [I] sont sujets à F(*x*). C'est le cas du verbe break 'briser' par exemple.

(51) The clock broke./James broke the clock.
(52) a. The breaking of the vase happened by itself.
b. James broke the vase on purpose.

3.3.2. **L'alternance inergative/transitive**

La fonction configurationnelle F(*x*) fournit un traitement de l'alternance inergative/transitive, supposant que la fonction F(*x*) dérive également la valeur [+A] pour la position Y en (53). Le verbe inergatif doit néanmoins être sous-spécifié pour le trait aspectuel terminatif [T].

(53) $[_v \ W \ [_v \ V_{[T]} \ \underline{Y}]]$

Nous prédisons que les verbes inergatifs téliques, soit les verbes qui sont aussi spécifiés [+T], ne peuvent être sujets à F($x$).

(54) a. The bomb exploded./*The bomb exploded the city.
b. La bombe a explosé./*La bombe a explosé la ville.

Nous prédisons aussi que les verbes inergatifs non téliques, tels que *briller* et *aboyer*, soit les verbes qui sont spécifiés [-T], ne peuvent pas non plus être sujets à F($x$), (55). Seuls les verbes qui sont sous-spécifiés pour [T], tels que *manger* et *parler* peuvent être sujets à F($x$), (56).

(55) a. The sun shines./*The sun shines the earth.
b. Le soleil brille./*Le soleil brille la terre.
c. The dog barks./*The dog barks the cat.
d. Le chien aboie./*Le chien aboie le chat.

(56) a. John spoke for/in five minutes.
Jean a parlé pendant/en cinq minutes.
b. John spoke the language *for/in five minutes.
Jean a parlé la langue *pendant/en cinq minutes.
c. John ate for/in five minutes.
Jean a mangé pendant/en cinq minutes.
d. John ate the apple *for/in five minutes
Jean a mangé la pomme *pendant/en cinq minutes.

### 3.4 **Résumé**

F($x$) couvre la formation de prédicats complexes sur la base de prédicats affixaux ainsi que la formation de prédicats complexes sur la base de prédicats libres. Ceci fournit une approche configurationnelle aux alternances ergative/transitive, inergative/transitive.

## 4. **Condition d'économie**

La fonction configurationnelle F($x$) permet de dériver des catégories étendues à partir de catégories élémentaires et de concert avec la sous-spécification lexicale rend compte du fait que certains

termes ont des valeurs catégorielles multiples et que certaines catégories forment des ensembles très limités. Cette fonction permet aussi de dériver des prédicats complexes à partir de prédicats élémentaires et de concert avec la sous-spécification sémantique rend compte des restrictions sur leur formation. Elle permet d'alléger les spécifications lexicales et contribue de manière significative à l'économie interne de la grammaire.

Références

Abney, S. 1987. *Noun phrase in its sentential aspect*. Thèse de doctorat, MIT, Cambridge, Mass.

Archangeli, Diana B. 1984. *Underspecification in Yawelmani phonology and morphology*. Thèse de doctorat, MIT, Cambridge, Mass.

Archangeli, Diana B. & Douglas Pulleyblank. 1994. *Grounded phonology*. Cambridge MA: MIT Press.

Chomsky, Noam. 1970. "Remarks on nominalization". In R. Jacobs & P. Rosenbaum (red.). *Readings in English transformational grammar*. Ginn, Walthman, Mass.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge Mass.: MIT Press.

Di Sciullo, Anna-Maria. 1995. "X-bar selection". In J. Roorick & L. Zaring (réd.). *Phrase structure and the lexicon*. Dordrecht: Kluwer, pp. 77-107.

Di Sciullo, Anna-Maria. 1996a. "Atomicity and Relatedness in Configurational Morphology". In A.M. Di Sciullo (réd.) *Configurations*. Sumerville : Cascadilla Press. pp. 17-40.

Di Sciullo, Anna-Maria. 1996b. "Modularity and $X^0$/XP asymmetry". *Linguistic Analysis* 26:3-28.

Di Sciullo, Anna-Maria & Elizabeth Klipple. 1994. "Modifying affixes", WECOL 23.

Di Sciullo, Anna-Maria & Edwin Williams. 1987. *On the definition of word*. Cambridge, Mass.: The MIT Press.

Emonds, Joe. 1985. *A unified theory of syntactic categories*. Foris, Dordrecht.

Grimshaw, Jane. 1990. *Argument structure*. Cambridge Mass,: The MIT Press.

Hale, Ken et Jay Keyser. 1993. "On argument structure and the lexical expression of syntactic relations". In K. Hale & J. Keyser (réd.) *The View from Building 20*. Cambridge, Mass.: The MIT Press.

Hendrick, Randal. 1991 "The morphosyntax of aspect". *Lingua* 85:171-210.

Houle, Jocelyne. en préparation. *Une analyse modulaire des constructions ergatives*. Thèse de doctorat, UQAM

Koopman, Hilda. 1995. "On verbs that fail to undergo V-second". *Linguistic Inquiry* 26:137-165.

Levin, Beth & Malka Rappaport Hovav. 1995. *Unaccusatives: at the syntax-lexical semantics interface*. Cambridge, Mass.: The MIT Press.

Lumsden, John. 1992. "Underspecification in grammar and natural gender". *Linguistic Inquiry* 23:469-487.

Rooryck, Johan. 1994. "On 0- and Å- Underspecification in syntax and phonology". Dans Harley et Collins (eds.) *The morphology-syntax connection*. MIT Working Papers in Linguistics, Volume 22. Cambridge: MIT.

Travis, Lisa. 1992. "Inner aspect and the structure of VP". *Cahiers de linguistique de l'UQAM* 1:130-144.

Van Vallin, R. 1987. "The unaccusative hypothesis versus lexical semantics: syntax vs semantic approaches to verb classification". In D. McDonough & E. Plunkett (réd.). *Proceedings of NELS 17*. GLSA. UMASS, Amherst.

Vendler, Zeno. 1967. *Linguistics and philosophy*. Ithaca: Cornell University Press.

Szabolsci, Anna. 1987. "Functional categories in noun phrase". In I. Kenesei (réd.). *Approaches to Hungarian* 2:42-63.

Williams, Edwin. 1981. "Argument structure and morphology". *The Linguistic Review* 1:81-114.

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 63-75.

# Le Schéma sémantique diachronique de Stern

*Evelyn Matheson-Styan*
Université Concordia[1]

**Introduction.** L'objet d'étude de la sémantique diachronique n'est pas parfaitement évident. Pour entreprendre une telle étude, d'après les linguistes qui travaillent dans ce domaine, cf. Geeraerts (1994), la première chose à faire est d'examiner attentivement les types de changements possibles dans les lexiques des langues. Deuxièmement, on détermine les causes des changements, et troisièmement, on étudie la transmission des changements au grand public. Les linguistes ne sont pas tous d'accord que ces étapes soient toutes nécessaires ou possibles, cf. Hock (1986:308), Lyons (1977:620), mais pendant les 150 dernières années, beaucoup de schémas ont été proposés pour la classification des changements lexicaux sémantiques. Quels sont les faits linguistiques ou les données à analyser?

Pour certains linguistes, la polysémie est une réflexion synchronique des changements sémantiques. Selon Guilbert (1975:64–65) la polysémie est un phénomène universel du lexique général d'une langue. Dans un bon dictionnaire ordinaire, l'entrée d'un mot polysémique présente tout un champ de sens différenciés les uns des autres et quelques sens reliés les uns aux autres. Pour le verbe *see*, on peut trouver au moins huit domaines de synonymes qui ne sont pas eux-mêmes des synonymes les uns des autres. De plus, on trouve un classement logique des sens regroupés en cinq domaines plus vastes (A-E). C'est cela qui corrobore la polysémie du verbe.

| | | | |
|---|---|---|---|
| A | *see* 1: | 'perceive visually' | *I saw her duck.* |
| B | *see* 2: | 'understand, realize, grasp' | *Do you see why?* |
| | *see* 3: | 'consider, judge, regard, view' | *You see things my way.* |
| C | *see* 4: | 'experience, go through' | *I've seen a happy life.* |
| D | *see* 5: | 'find out, check, ascertain' | *Let's see what's on TV.* |
| | *see* 6: | 'meet, visit, consult' | *See Jo for a shampoo.* |
| E | *see* 7: | 'make sure, attend to, look after' | *She'll see to it.* |
| | *see* 8: | 'escort, accompany, go with' | *I'll see you to the door.* |

[1]Je dois mille remerciements aux lecteurs francophones suivants: Serge Alalouf, Julie Auger, Michel Bertrand, Maryse Dansereau, et Yvan Rose. Je suis bien sûr l'unique responsable des erreurs restantes.

En A, on trouve le sens le plus fondamental de 'percevoir par les yeux'. En B, le référent du sujet a des idées sur quelque chose. C concerne l'expérience du référent en entier. En D, on cherche des renseignements. En E, on s'assure de quelque chose. [Cette analyse est fondée sur l'étude synchronique détaillée du verbe *see* par Alm-Arvius (1993); voir aussi une autre étude par Sweetser (1990).]

On peut prendre les définitions (données par les synonymes partiels) comme indiquant différentes acceptions du verbe *see*, tout en mettant en doute que cet exposé, fait rétrospectivement, représente le lexique de l'anglais contemporain, surtout chez un locuteur compétent. Alors l'organisation logique ne reflète pas forcément l'évolution du verbe *see* ou des événements explicites de changement sémantique. Car chaque événement de changement est l'effort accompli par un individu (le sujet parlant) et peut être daté (en théorie, mais à peine en pratique). Lyons (1977:620–621) a observé que ce phénomène est une réflexion de variation non seulement dans le temps mais aussi dans l'utilisation (dialectale) par le grand public (en groupes) pour divers besoins à la base sociaux.

Pour le linguiste, il ne s'agit pas non plus de chercher la véritable histoire du mot. Aussi intéressante et unique qu'elle soit, on ne peut l'apprendre que rarement. L'histoire des mots est aussi difficile pour les lexicographes, comme créateurs des dictionnaires synchroniques ou diachroniques. Au mieux, les linguistes ou les meilleurs dictionnaires donnent l'histoire des mots seulement dans ses lignes générales. Par exemple, on sait en gros que le verbe *travel* en anglais moderne est dérivé étymologiquement de l'ancien français *travail*, *travailler*, ce qui suggère que l'activité de voyager au moyen âge était ardue. Bien sûr, l'activité correspondante de nos jours ne se passe pas toujours sans effort, et la signification du verbe *travel* varie selon la situation donnée, d'après Gustaf Stern (1931). Ce genre de "changement sémantique" qu'il appelle "substitution" [de référent] est dû à des causes extérieures au système linguistique. C'est la première des sept classes de changements sémantiques que Stern a proposées. L'objet de Stern a été non seulement une taxonomie des changements sémantiques qui sont possibles dans les langues, mais aussi une explication de leurs causes. Son traitement de la transmission des changements est très bref, mais on peut s'attendre à ce que l'explication de la diffusion des néologismes suive les même principes que les changements (originaux), mais à un plus haut degré. Ce sujet mérite peut-être un traité sociologique à part.

Cet article ne présente qu'un bref résumé du schéma diachronique de Stern; j'ai tenté d'appliquer sa classification à de nou-

velles polysémisations en anglais, tout en me demandant quel profit on peut tirer du système pour l'analyse linguistique. Je considère quelques exemples de l'application des idées de Stern en deux parties: (1) le schéma sémantique de Stern, (2) les causes d'après Stern des changements de sens lexicaux. Je terminerai avec (3) des remarques concernant les qualités intrinsèques et les difficultés que j'ai éprouvées dans l'application de son système.

**1. Classification de changements de sens lexicaux selon Stern.** Beaucoup de linguistes (du passé) ont admiré en partie l'ouvrage de Stern (1931). Par exemple, Ullmann (1967:246-249) l'a trouvé clair, documenté, et fondé sur un empirisme robuste. Ullmann a admiré la méthode de travail de Stern et il a même considéré le système de Stern comme étant aussi complet que le sien. Guiraud (1979:45) a également trouvé les deux systèmes [de Stern et d'Ullmann] comparables, mais il a noté en plus que l'ouvrage "de G. Stern constitue l'inventaire le plus complet des changements de sens considérés du point de vue causal et fonctionnel ... et classés en même temps selon un cadre sémiologique ... " Stern base son schéma sur une étude des mécanismes cognitifs. Néanmoins sa présentation du système (avec beaucoup d'exemples tirés de l'histoire de l'anglais, entre autres langues) et son argumentation approfondie ne suffisent pas à convaincre les linguistes des générations suivantes d'adopter ou même de remanier son schéma; cf. Traugott (1985), Lyons (1977). On peut se demander pourquoi.

Cependant sous certains rapports il me semble que Stern a été en avance sur son temps. Étant donné le fort intérêt actuel pour la langue et la connaissance, son étude me semble assez moderne, même si Stern lui-même (1931:16) ne serait pas d'accord: la base psychologique dont il s'est servi commence à dater (après 65 ans). Néanmoins, en dépit des hypothèses pénétrantes qui sont disponibles aujourd'hui, surtout en ce qui concerne la localisation des activités cérébrales, les faits qu'on peut considérer établis sur l'organisation des systèmes de connaissance et de la langue semblent peu nombreux. L'étude de Stern reste assez profonde et ses notions concernant les aspects cognitifs et fonctionnels de la langue me semblent raisonnables. Stern considère les changements de sens du point de vue tant pragmatique que sémantique (sans préciser cette distinction). Son livre traite des problèmes généraux de définition de la signification lexicale de même que des problèmes de production et de compréhension du discours.

Comme partie intégrante de la classification des changements lexicaux sémantiques, Stern donne des critères nets mais généraux. Tout d'abord, il distingue les changements extra-linguistiques des changements linguistiques. Cf. A et B. Puis il propose trois groupes de changements linguistiques. Le résumé de son schéma (Stern 1931:175) est le suivant:

A. External Causes: Class **1 Substitution**

B. Linguistic Causes:
- I. Shift of Verbal Relation
  - a. Class **2 Analogy**
  - b. Class **3 Shortening**
- II. Shift of Referential Relation
  - a. Class **4 Nomination**
  - b. Class **5 Transfer**
- III. Shift of Subjective Relation
  - a. Class **6 Permutation**
  - b. Class **7 Adequation**

Avant de présenter les sept classes d'une façon plus détaillée, je voudrais examiner la source de ces classes et les hypothèses de base. D'où vient cette classification de changements de sens? J'ai déjà mentionné les méthodes de travail empiriques de Stern. Après avoir recueilli des exemples de mots simples, il les a analysés selon des traits qui lui semblaient significatifs. Enfin il les a classés en sept groupes. Pour s'assurer que les sept groupes de changements n'étaient pas seulement des "regroupements statistiques", Stern cherchait à définir les classes selon des buts linguistiques. Il était très heureux de pouvoir les caractériser en fonction du modèle sémiologique. Pour lui, c'était une vérification de sa classification. Les sept classes de Stern sont déterminées par les interrrelations entre les trois points du triangle sémantique proposé par Ogden et Richards (1923). [Le développement de cette hypothèse est bien expliqué par Ullmann (1967) et résumé par Lyons (1977:ch. 4).] Examinons maintenant la définition de la signification selon Stern.

L'interprétation donnée du triangle Ogden-Richards ci-dessous est propre à Stern. Je présente le système de Stern en format linéaire. Les numéros 1, 2, 3 désignent les trois points du triangle, avec les étiquettes traduites du texte de Stern, sauf en ce qui concerne le raccourcissement de "mental content" à *contenu.* Dans son texte Stern emploie aussi le terme *concept* pour désigner ce qui s'interpose entre l'expression et le ou les référents.

1 Expression <--> 2 Contenu <--> 3 Référent(s)

D'après ce modèle, le sens d'une expression est déterminé par trois éléments dont les interrelations sont causales. Selon Stern (1931:38) (1) l'expression exprime (2) le contenu (sens, pensée) et nomme ou dénote (3) le référent. Il me semble bon de souligner les sources et les caractéristiques des éléments impliqués par Stern: 1 = les éléments du système linguistique, les signifiants; 2 = le sens, "meaning" au sens étroit, ce qui est subjectif pour Stern, chez le locuteur, basé sur l'appréhension du référent; 3 = ce dont on parle, le ou les référents, toujours objectifs, puisqu'ils peuvent être les mêmes pour plusieurs interlocuteurs. Bien sûr, l'appréhension du référent (la conception) diffère selon le sujet parlant. (Il y a pour toute la communauté linguistique tout un champ de conceptions, parmi lesquelles se trouve le concept. C'est l'élément noétique, une conception "intersubjective" que partagent tous les sujets parlant une même langue qui connaissent l'expression donnée. Le contenu est alors une conception que tout le monde accepte). Les mots et leurs contenus sont étroitement reliés plus ou moins automatiquement dans l'esprit du locuteur (Stern 1931:112). Il suppose que le mot, quoique polysémique, ne peut être associé qu'à une seule conception tant pour le locuteur que pour l'auditeur. De plus, pour Stern (1931:105), la langue est le moyen d'expression aussi bien que la méthode d'analyser les sensations et le contenu de la pensée.

Le changement de sens d'un mot résulte en pratique de l'ajustement dans l'esprit du locuteur de la relation à une situation donnée dans le contexte d'utilisation (Stern 1931:116). Un changement de sens a lieu au moment où un locuteur forme et attache une nouvelle conception à une expression déjà existante. Les changements de sens ont pour résultat (1) de dénoter un ou plusieurs référents nouveaux, ou (2) d'exprimer une nouvelle façon d'appréhender ou de percevoir le(s) référent(s). Enfin Stern base son schéma sémantique diachronique sur les déplacements parmi les éléments impliqués dans l'analyse tripartite. Stern nous fait mieux comprendre les distinctions entre les sept classes en citant des exemples qu'il considère bien caractéristiques. De plus, il décrit ces exemples en termes de linguistique traditionnelle, quoique ses classes ne se limitent pas aux catégories traditionnelles. Ces sept classes seront maintenant présentées d'une façon plus détaillée et illustrées avec des exemples de Stern (1931) et de Guiraud (1979) et aussi des nouvelles polysémisations de l'anglais. J'ai tiré mes exemples de périodiques récents, p. ex.,

*The New York Times, Newsweek, The New Yorker*, les bandes illustrées de *The Gazette* (Montréal), etc.

Dans les cas des **changements externes**, il s'agit d'un changement de référent sans changement de nom. En citant l'exemple de *travel* ci-dessus, je voulais illustrer ce genre de changement qui se nomme "Substitution" (classe **1**). On en distingue trois types. Une substitution peut être: (**1a**) objective, p. ex., *travel, ship, plume, livre*, etc. (**1b**) cognitive, p. ex., *atome* (n'étant plus considéré comme l'élément physique irréductible) ou (**1c**) subjective, p. ex., *gay* ('homosexuel', terme péjoratif > terme neutre ou plus objectif). Un cas de substitution est classé selon (a) que le changement se produit réellement dans le monde décrit, (b) qu'il s'agit de la définition ou des critères de catégorisation, ou (c) qu'il dépend tout simplement de l'attitude du locuteur envers le référent. Cette première classe de "changements sémantiques" est due au "conservatisme linguistique" (Ullmann 1967:220). Les sujets parlants continuent d'employer le même mot pour dénoter ce qu'ils considèrent être la même activité (*travel*), le même type de véhicule marin (*ship*), etc.

Des **changements linguistiques** (en ce qui concerne le système linguistique), Stern distingue 3 types principaux tels que présentés ci-dessous. Ce sont les déplacements (I) de la relation verbale, (II) de la relation significative, et (III) de la relation subjective. Dans la première catégorie linguistique, il est question tout simplement d'un déplacement de la relation verbale (du nom) sans changement de sens ou de référent [1 Expression <--> 2 Contenu <--> Référent(s)]. Le sujet parlant crée un nouveau signifiant pour désigner un concept donné, ignorant ou non le signifiant standard. Il y en a deux classes: **Analogy** (classe **2**) et **Shortening** (classe **3**). Les cas d'analogie dépendent de la connaissance de la grammaire de la part du locuteur, car le cas typique est un changement de forme du mot d'une certaine classe lexicale par analogie. Les exemples de Stern sont divisés en trois sous-classes: (**2a**) reformations morphologiques: dérivation, composition, flexions, p. ex., (du langage des enfants) *went = goed;* (**2b**) corrélative: calques, p. ex., *réaliser* (FRAN.) = *realize* (ANG.); et (**2c**) phonétique: contagion, étymologie populaire, p. ex., *prostrate / prostate cancer.* Quant à la deuxième classe de changements verbaux (Shortening), il y a deux sous-classes de raccourcissements: (**3a**) troncation; p. ex., *auto(mobile) parts, ute < utility vehicle, mini-ute;* et (**3b**) ellipse: p. ex., *des (pommes) frites, la (ville) capitale; la (église) cathédrale.* Les formes des exemples ci-dessus sont dérivées par l'omission des éléments donnés entre parenthèses. On peut ajouter à cette classe la siglaison

et l'acronymie, p. ex., *CEO* < chief executive officer, *AIDS* < *Acquired Immune Deficiency Syndrome*. *E-book, e-memo, e-letter* sont tous des documents électroniques.

Quant au déplacement de la relation significative ou conceptuelle [1 Expression <--> 2 Contenu <--> 3 Référent(s)], il y a deux sous-classes qui se distinguent selon l'intentionnalité du locuteur. Pour un cas de **Nomination** (classe **4**), le locuteur donne un nouveau nom à un concept intentionnellement, tandis que pour le changement que Stern nomme **Transfer** (classe **5**), le transfert est non intentionnel. On ne peut guère distinguer les candidats de ces deux classes (4 et 5) sans connaître le contexte immédiat des énoncés en cause. Car chaque énoncé est l'expression de l'état momentané de l'esprit et du but d'un locuteur, selon Stern (1931:162). Mais en étudiant ses exemples, on peut proposer au moins le classement des néologismes à analyser.

De la Nomination, Stern distingue les trois sous-classes suivantes: (**4a**) composition et dérivation, p. ex., *e-memo, e-letter, e-book* ; dans *the tony auction of Jackie O*, l'adjectif *tony* < *(high) tone + y; sportswriterese* < *sports + write + -er + -ese; dissing* < *dis-(X) + -ing*. L'amalgame est aussi un processus productif, p. ex., *affluenza* < *afflu(ent) + (influ)enza; anticipointment* < *antici(pation) + (disap)pointment*. (**4b**) création de mots: les onomatopées, p. ex., *whoosh*. On doit se demander comment on peut considérer des néologismes (au sens plus étroit de la création de mots *ex nihilo*, si une telle chose est possible) comme étant un "changement" sémantique. Je souligne une création (tirée de la bande illustrée *Doonesbury*) qui est dérivée d'un emprunt à l'argot interne.

*Then after that we just geeked out on the Beta games I bought.*

(**4c**) transfert intentionnel, y compris les figures telles que la métaphore stylistique. L'exemple qui suit contient un verbe (en souligné) qui est employé (on peut supposer d'une façon intentionnelle) pour désigner les référents ou plutôt leurs attributs au sens figuré. Autrement quand le choix d'une expression dite "métaphorique" est-il fait intentionnellement?

*It's hard to read a little kid, Mike.*

L'analyse de Stern nous éclaire sur les métaphores stylistiques, telles qu'élaborées par Tournier (1985:ch. 6). Bref, ici des aspects de la portée du prédicat *read* s'appliquent aux référents du sujet sous-entendu et de l'objet grammatical. (Le verbe *read* est-il un synonyme de *see* 2 et 3?)

En contraste avec le prédicat implicite du transfert intentionnel, le transfert (non intentionnel) est explicite. Les exemples pour les trois sous-classes de la classe **5**, nommée **Transfer**, sont tirés du livre de Guiraud (1979:48): (**5a**) identité d'apparence, p. ex., *sheet of paper, feuille de papier;* (**5b**) identité de fonction, p. ex., *a riverbed, lit d'une rivière* ; (**5c**) identité de situation, p. ex., *foot of a mountain, pied d'une montagne.*

Quant au déplacement de la relation subjective entre le mot et les sujets parlants [1 Expression <--> 2 Contenu <--> Référent(s)], il y a deux sous-classes, **Permutation** (classe **6**) et **Adequation** (classe **7**). La permutation constitue un changement dans l'idée que le sujet parlant se fait du référent. On peut voir des exemples familiers des trois sous-classes de permutations: (**6a**) métonymie, p. ex., *un marbre* (la matière pour l'objet); *du bordeaux* (la région pour le produit); (**6b**) synecdoque: p. ex., *une voile* (la partie pour le tout); (**6c**) permutation de 2 objets: p. ex., *bead:* 'prayer' > 'bead'. Pour Stern, l'adéquation (classe 7) est un "mode particulier de la classe (6) relatif à un nouveau caractère du référent", p. ex., *horn* (ANG.) 'corne d'un animal ' > 'cor (de chasse)' > 'cornet à piston'.

D'après Stern (1931:168-169) l'exemple de *horn* représente tout un enchaînement de changements: une substitution, une adéquation, et une deuxième substitution. Par ailleurs, Stern lui-même combine le type de changement 7 à tous les autres. Donc les changements particuliers que je viens de citer semblent pouvoir appartenir à plusieurs types en même temps. Ne peut-on pas dire avec Guiraud (1979:71) que "la plupart des changements de sens sont le résultat de procès multiples... "? C'est bien évident. Dans l'exemple

*I'm feeling the skizzy twinges of having subdivided myself.*

l'auteur nous laisse comprendre le sens du néologisme *skizzy* en en donnant une définition: "having subdivided myself. " *Skizzy* est dérivé du terme technique *schizophrenic,* dont l'auteur a découpé la racine et modifié l'orthographe, et à laquelle il a ajouté le suffixe *-y* pour former l'adjectif. Alors ce mot est une création qui appartient aux classes (3a) et (4a), ou la troncation et la dérivation.

Or il semble que les changements d'un certain type entraîneraient des changements d'un autre type. Par exemple, des syntagmes nominaux qui sont longs, compliqués, mais en vogue, sont de bons sujets pour le raccourcissement. Ce que tout le monde connaît comme *mad-cow disease* se nomme aussi *Bovine Spongiform Encephalopathy* > *BSE.* Tout d'abord, il y a la création des syntagmes par un processus de créativité lexicale que Guilbert

(1975:101) appelle "la néologie syntagmatique". En effet, c'est de la composition syntaxique. Un syntagme, parfois toute une phrase, peut fonctionner comme un attribut ou comme un prédicat.

> C'est un *"m'as-tu-vu?"*.
> the *have-* and *have-not* nations; the *haves* and the *have-nots*.
> Nevermind the *gloom-and-doomisms* of modern life.
> I'm just an urban country *wannabe;* wannabe presidents / stars.

Ces constructions sont-elles intéressantes ou tout simplement des "fluctuations" (Stern 1931:163) ou des "glissements d'emploi" (Tournier 1985:45) qui entraîneront enfin leur emploi usuel? Étant donné la nature variable et flexible du lexique, on ne peut prédire si ces compositions syntaxiques subiront un changement de troncation plus tard. Ces exemples semblent renforcer la notion que le mécanisme de "réinterprétation" de Hock (1986), qui s'applique à beaucoup de types de changements linguistiques, s'applique surtout aux changements sémantiques. Comme on peut dire pour la phrase entre guillemets dans *C'est un "m'as-tu-vu?"*, l'expression *wannabe* est peut-être dérivée du texte (réinterprété) d'un acte de langage à la première personne du singulier. Le locuteur semble exprimer une attitude particulière envers le référent, tout en lui attribuant le comportement de quelqu'un qui a l'ambition de briller d'une manière ou d'une autre.

**2. Causes des changements lexicaux.** La façon dont Stern aborde le problème tout en considérant les questions linguistiques, psychologiques, et en même temps logiques, me semble non seulement courageuse mais aussi appropriée, peut-être ce qu'il faut faire pour étudier les changements de la langue comme partie intégrante de la vie humaine. D'après Slobin (1975) la structure de la langue même peut être déterminée par des contraintes diverses: de la perception, de la mémoire, de la connaissance, et même du développement de ces capacités dans des situations sociales pendant la jeunesse.

Quoique les causes ultérieures de changement lexical puissent être économiques, historiques, psychologiques, sociales, scientifiques, techniques, industrielles, etc., pour généraliser, Stern prétend qu'il n'y a en définitive que quatre causes. D'après lui, ces causes définitives résultent des fonctions langagières, étant dues aux efforts du locuteur pour symboliser justement ses perceptions des choses, pour bien s'exprimer, pour réussir à communiquer avec les autres, et pour exercer une influence sur leur pensée. Il les nomme

donc selon les fins cognitives, communicatives, expressives, et affectives. Les causes définitives des changements de sens lexicaux sont les fonctions de la parole humaine. Selon Stern (1931: ch. 2) ces fonctions sont:

(i) "communicative" (l'importance de dire quelque chose à un ou plusieurs auditeurs)
(ii) "symbolic" (le choix des signes logiques pour bien symboliser le référent ou les référents)
(iii) "expressive" (l'expression des émotions, pensées, sentiments du sujet parlant)
(iv) "purposive" (l'objectif de la communication).

Stern a remarqué que, même si elle était complète, cette liste de fonctions ne pouvait servir au linguiste comme base de la classification des changements de sens: une cause donnée peut déterminer un changement de type 1 à 7 ou plusieurs changements différents à la fois. Il observe aussi qu'il y a deux tendances opposées dans l'usage d'un système linguistique: la productivité et la stabilité. Pour être compris, les besoins de symboliser et de communiquer forcent le locuteur à employer les mots conventionnellement et selon le système standard des règles linguistiques. Lorsqu'un locuteur emploie le système linguistique pour des fins expressives, il peut l'exploiter d'une manière créatrice. La fonction affective nécessite une attention aux besoins des trois autres fonctions.

Pour la création des néologismes, il y a une certaine évidence de l'importance de l'utilisation des sources préexistant dans le lexique. Dans l'étude statistique qu'a faite Algeo (1980) des néologismes ajoutés à un dictionnaire anglais au cours d'une récente décennie, il a trouvé que les deux tiers des 1 000 mots choisis par sondage (68,7%) étaient de type composé, y compris l'amalgame, la composition, et la dérivation. J'ai cité ci-dessus des exemples de ces types de formation. Les autres types importants de formation sont: la conversion (14,2%), le raccourcissement, surtout les acronymes ou les abréviations (9,7%), et l'emprunt (6,9%). Évidemment la grande majorité des changements lexicaux sont des cas de recyclage des corrélations familières forme-sens, que Algeo (1980) ne tente pas d'expliquer. Mais on peut s'attendre à ce que les causes données par Stern puissent suggérer une explication à cet égard.

**3. Commentaire.** Les aspects du système sémantique diachronique de Stern que je trouve les plus intéressants ont trait à l'attention qu'il accorde au sujet parlant comme catalyseur des

changements lexicaux. D'après son analyse de la signification, avant tout de l'interprétation sémantique, il entrevoit la complexité des interrelations causales en question. Mais ce système compliqué peut devenir en même temps la force et la faiblesse de l'analyse. À tout prendre, j'admire les mécanismes que Stern a proposés mais je vois les problèmes que lui suggère leur application.

Quant à l'analyse tripartite, un léger problème se pose. Stern prétend que pour tout changement de sens, deux des trois relations restent stables tandis qu'une troisième se déplace. Si les relations sont bien causales, un premier déplacement au point 1 en entraînerait un deuxième au point 2. Autrement dit, une addition au nom entraînerait deux autres déplacements. Alors, comment peut-on garder certaines relations fixes tout en en changeant d'autres? Sa liste de composantes du système linguistique peut être complète, mais tout comme Stern a remarqué en ce qui concerne les causes de changements, il n'est pas évident que le linguiste puisse se servir des composantes du modèle comme base de la classification des changements de sens; les interrelations peuvent être si bien ajustées que pour toute interprétation, toutes les composantes sont impliquées à la fois, qu'il y ait un changement ou non. Le problème de base est que le système peut être si complexe qu'il est irréductible en ses parties composantes; chaque soi-disant "élément" joue un rôle si essentiel dans le fonctionnement du système entier, qu'on ne peut omettre aucune de ces composantes sans risquer la confusion ou la non interprétation. N'est-il pas possible que les matériaux de construction pour toutes les composantes soient remplacés continuellement? À la suite des idées de Stern en ce qui concerne la flexibilité de sens, le sens d'un mot peut facilement se changer avec des appréhensions variables des locuteurs. Il suggère que le sens du mot *bird* varie selon le type d'oiseau auquel il est appliqué. Si quelqu'un dit, *there's a bird*, le sens diffère pour un moineau, une corneille, ou un rouge-gorge (Stern 1931:40).

Un problème semblable concerne les changements de type 1 du groupe A, dus aux causes externes. Seul le changement (1a) qui concerne les changements dits "objectifs" peut être dû exclusivement aux causes externes. Et la notion d'une telle classe de "changements sémantiques" est vraiment inconcevable, s'il n'y a pas de changement (du sens des mots résultant de l'ajustement de la relation) dans l'esprit du locuteur à la situation donnée dans le contexte d'utilisation (Stern 1931:116). Le référent peut changer matériellement mais, sans adéquation, il ne s'agit pas de changement de sens. Par contre, les changements de type (1b) et (1c) sont tous les deux de véritables

changements de sens, résultant de l'ajustement de la relation à une situation dans l'esprit du locuteur. Cette proposition, si elle s'avère correcte, nous laisse mieux comprendre la complexité des relations qui existent entre le langage et la réalité extra-linguistique.

Stern applique beaucoup d'attention à l'étude du locuteur, mais il est possible qu'il ne s'intéresse pas suffisamment aux distinctions linguistiques. Ce problème est grave du point de vue de l'analyse linguistique ordinaire. Sauf pour les classes 1 et 2, qui résultent de déplacements de la relation verbale, les classes de Stern ne tiennent pas compte des critères formels. Par exemple, la nomination (classe 5) se manifeste sous des formes très variées, selon que la nouveauté lexicale concerne la création des mots par construction morphologique, motivation phonique, ou métasémie. Il manque les généralisations linguistiques comme on doit s'y attendre, étant donné le fort intérêt des linguistes pour la forme, et surtout pour la corrélation qui peut exister entre la forme et le sens des items lexicaux. Aujourd'hui, comme autrefois, les changements de sens seraient classés dans des catégories telles que les changements morpho-sémantiques (la composition, l'affixation, les onomatopées); les changements sémantiques (la conversion, la métasémie); les changements morphologiques (p. ex., la troncation, la siglaison) et les emprunts; cf. Guilbert (1975); Tournier (1985).

Malgré l'importance numérique des formations dérivées d'éléments existant déjà dans la langue, il n'est pas toujours possible de dire quelle acception du mot a été utilisée comme matrice. À titre d'exemple citons le verbe polysémique *go* dans le dialogue suivant:

*She goes: "I am going to go to Paris." And then he goes: "Oh no, you're not going!"*

Certaines formes du même lexème *go*, qui signifie fondamentalement 'aller d'un endroit à un autre' ou 'quitter (un endroit, une personne)', peuvent être employées pour exprimer 'dire' surtout comme à tour de rôle et la fonction "auxiliaire" *be going to* + inf. qui veut dire que le locuteur s'attend à ce que le référent du sujet fasse quelque chose. L'utilisation de *go* pour introduire une citation, une exclamation, etc. constitue-t-elle en fait une nouvelle polysémisation du verbe *go*, ou est-elle reliée au nom *go* dans l'expression *to have a go*, 'faire quelque chose à tour de rôle'? Il est pratiquement impossible de déterminer quelle acception précise de *go* en a été la source. Au lieu de *He went : "X ..."* ou *He was going: "X ..."* on aurait pu dire *He was like: "X ..."* L'emploi des prédicats *be like* et *go*, qui se fait au présent ou au passé,

semble ajouter une réelle qualité dramatique à la narration. Ils ne sont pas suivis de citations indirectes.

* *She's like / she goes that she will leave.*

De tels faits morpho-syntaxiques semblent évidents. Cependant pour l'analyse sémantique diachronique, faite rétrospectivement, les linguistes ne peuvent que deviner les rapports entre les acceptions des expressions données, comme font souvent les lexicographes, qui nous offrent quand même de très bons outils de travail.

## BIBLIOGRAPHIE


Algeo, John. 1980. "Where do all the new words come from?". *American speech* 55:264–277.

Alm-Arvius, Christina. 1993. *The English verb see: a study of multiple meaning*. Göteborg, Acta Universitatis Gothoburgensis.

Geeraerts, Dirk. 1994. Articles sur "Historical semantics", "Semantic laws", etc. In R. E. Asher, ed. *The Enclyclopedia of language and linguistics*. Oxford: Pergamon Press.

Guilbert, Louis. 1975. *La créativité lexicale*. Paris: Larousse.

Guiraud, Pierre. 1979. *La sémantique*. 9e éd. Paris: P.U.F.

Hock, Hans H. 1986. *Principles of historical linguistics*. Berlin: Mouton de Gruyter.

Lyons, John. 1977. *Semantics*. 2 v. Cambridge University Press.

Ogden, C. K. & I. A. Richards. 1923. *The meaning of meaning*. London: Routledge & Kegan Paul.

Slobin, D. I. 1975. "The more it changes: on understanding language by watching it move through time". In *Papers and reports on child language development* (Department of Linguistics, Stanford University), no. 10, pp. 1–30.

Stern, Gustaf. 1931. *Meaning and change of meaning, with special reference to the English language*. Bloomington: Indiana University Press. (3d printing 1968.)

Sweetser, Eve. 1990. *From etymology to pragmatics: metaphorical and cultural aspects of semantic structure*. Bloomington: Indiana University Press.

Tournier, Jean. 1985. *Introduction descriptive à la lexicogénétique de l'anglais contemporain*. Genève: Champion-Slatkine.

Traugott, Elizabeth C. 1985. "On regularity in semantic change". *Journal of literary semantics* 14:155–173.

Ullmann, Stephen. 1967. *The principles of semantics*. 2nd ed. Oxford: B. Blackwell.

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 77-88.

# Les catégories sémantiques «événement» et «action»

*Anne Rochette*
Université du Québec à Montréal

## 1. Introduction

Les travaux de Grimshaw (1979, 1981) et de Pesetsky (1982) ont démontré la nécessité de recourir à un mécanisme de sélection sémantique et ont ainsi ouvert la voie à un programme de recherche qui s'inscrit à l'interface de la sémantique lexicale et de la syntaxe.

Ce programme de recherche, quoique prometteur, se heurte malheureusement à l'épineux problème de l'identification et de la caractérisation des catégories sémantiques. En effet, la justification de l'existence d'une catégorie sémantique particulière dépend souvent de considérations à la fois syntaxiques et sémantiques, ce qui amène parfois une certaine forme de circularité dans l'argumentation. De plus, il est essentiel dans ce type de recherche de s'astreindre à questionner la validité de chacune des catégories postulées afin de ne pas multiplier inutilement le nombre de catégories, cette démarche étant d'autant plus importante que ces catégories sont considérées comme des primitifs sémantiques.

Dans des recherches antérieures portant sur la réalisation et la distribution des complétives ainsi que sur la distribution des adverbes (Rochette 1988, 1991), j'ai proposé de distinguer les catégories sémantiques *proposition*, *événement* et *action*. Des recherches plus récentes sur les restrictions de sélection imposées par les verbes aspectuels m'ont également amenée à proposer l'existence de la catégorie sémantique *processus* qui caractérise les complétives sélectionnées par ces verbes (Rochette 1993).

Dans cet article, j'aimerais questionner la validité de l'opposition postulée entre les catégories sémantiques *événement* et *action* et, plus particulièrement, remettre en question l'existence de la catégorie *action*. Dans un premier temps, je passerai en revue les différents arguments qui ont conduit à la postulation de cette catégorie pour ensuite montrer que ces arguments ne sont pas vraiment décisifs et qu'il est possible de rendre compte des différentes propriétés examinées tout en n'ayant recours qu'à la seule catégorie *événement*.

## 2. Justification de l'opposition *événement / action*

Dans Rochette (1988), j'ai argumenté que la distribution et les propriétés des complétives infinitives et tensées d'une langue comme le français peuvent être expliquées en termes d'une opposition entre trois catégories sémantiques : les catégories *proposition*, *événement* et *action*.

### 2.1. La distribution des complétives

Les verbes qui introduisent des complétives se regroupent en trois grandes classes, selon le ou les types de complétives qu'ils permettent. Ainsi les verbes épistémiques, d'assertion et de connaissance peuvent introduire des complétives tensées à l'indicatif ainsi que des complétives infinitives (1), les verbes de volonté, de désir, de permission et de commande permettent des complétives tensées au subjonctif ainsi que des infinitives (2), alors que les verbes aspectuels, les modaux et les verbes de mouvement ne permettent que des complétives infinitives (3).

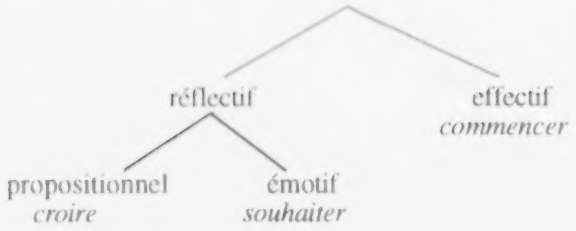
(1) a. Jean croit avoir compris le problème.
b. Jean croit qu'il comprend le problème.
c. *Jean croit qu'il comprenne le problème.

(2) a. Jean souhaite partir.
b. Jean souhaite que Marie parte demain.
c. *Jean souhaite que Marie partira demain.

(3) a. Jean commence à écrire son livre.
b. *Jean commence qu'il écrit son livre.
c. *Jean commence qu'il écrive son livre.

Afin de rendre compte de la sélection des divers types de complétives ainsi que de leurs nombreuses différences de comportement syntaxique, j'ai proposé que celles-ci correspondent à des catégories sémantiques distinctes sélectionnées par des classes sémantiques de prédicats également distinctes. J'ai adopté la classification sémantique des verbes matrices proposée par Long (1974) et illustrée ici en (4).

(4)




Selon Long, les prédicats réflectifs expriment les jugements d'un sujet (humain) par rapport à une proposition ou un événement alors que les prédicats effectifs décrivent plutôt la relation d'un sujet, qu'elle soit causale, potentielle ou autre, par rapport à la performance d'une action. Les prédicats réflectifs comprennent deux grandes classes : les prédicats propositionnels qui expriment des jugements de valeur de vérité par rapport à une proposition et les prédicats émotifs qui expriment des jugements d'ordre personnel par rapport à un événement.

J'ai proposé que chaque classe sémantique de verbes sélectionne une catégorie sémantique distincte dont la réalisation syntaxique serait également distincte. Ainsi, les prédicats effectifs sélectionnent la catégorie sémantique *action* dont une des réalisations canoniques structurales est une projection de la catégorie V, ce qui rend compte de l'absence de compléments tensés avec ces prédicats. Les prédicats émotifs sélectionnent la catégorie sémantique *événement* dont la réalisation canonique structurale est une projection de la catégorie Infl, ce qui permet le choix du mode subjonctif ou de l'infinitif.[1] Les prédicats propositionnels sélectionnent la catégorie *proposition* dont la réalisation canonique structurale est une projection de la catégorie Comp, qui permet soit une complétive indicative ou une complétive infinitive.[2]

---

[1] Compte tenu des développements ultérieurs dans la recherche sur les catégories fonctionnelles, la réalisation canonique structurale de la catégorie sémantique *événement* serait probablement une projection de la catégorie Mode (*Mood*).

[2] Un quatrième groupe de verbes qui semble tenir à la fois des verbes émotifs et des verbes propositionels a aussi été distingué : il s'agit des verbes émotifs factifs (e.g. *regretter*). Suivant Melvold (1986), j'ai proposé qu'ils sélectionnent la catégorie *événement défini*.

Ce premier argument soutenant la distinction de trois catégories sémantiques est donc essentiellement basé sur la distribution de trois types de complétives par rapport à trois classes de prédicats matrices. La caractérisation des complétives des verbes propositionnels et émotifs en termes des catégories *proposition* et *événement* semble nettement justifiée, le statut de ces catégories en tant que primitifs sémantiques étant relativement clairement établi dans les travaux en sémantique. Cependant, il est moins clair que la catégorie *action* puisse se prévaloir de ce statut de primitif et il convient de se demander si les complétives des verbes effectifs ne relèvent pas elles aussi de la catégorie *événement*.

### 2.2. Les propriétés des complétives

L'argument basé sur la distribution des complétives reçoit toutefois une justification supplémentaire dans le fait qu'un certain nombre de propriétés sémantiques et syntaxiques distinguent les structures dans lesquelles apparaissent les prédicats effectifs des structures qui impliquent des prédicats émotifs ou propositionnels. J'aimerais maintenant examiner quelques-unes de ces propriétés.

#### 2.2.1. Les propriétés sémantiques des complétives

Sur le plan de l'interprétation sémantique, les prédicats effectifs partagent la propriété de ne pas pouvoir constituer un événement autonome et ont plutôt comme fonction d'apporter une modification d'ordre aspectuel ou implicationnel à l'événement décrit par le complément infinitif. Ainsi, l'interprétation des phrases données en (5) ne comporte pas deux événements distincts, dont chacun serait constitué de l'un des deux prédicats présents dans la phrase, mais comporte plutôt un seul événement constitué d'un complexe prédicatif qui englobe les deux prédicats. L'impossibilité d'utiliser des adverbes temporaux dénotant des moments distincts démontre également qu'un seul événement est en cause avec les prédicats effectifs.

(5) a. Jean commence à lire ce livre.
b. Jean ose lire ce livre.

(6) a. *Hier, Jean a commencé à lire ce livre aujourd'hui.
b. *Demain, Jean osera avoir lu ce livre aujourd'hui.

Par contre, l'interprétation de phrases qui impliquent un prédicat émotif comme *souhaiter* ou un prédicat propositionnel comme *constater* nous montre que ces prédicats constituent des événements autonomes par rapport à ceux dénotés par leur argument interne. Ceci est également confirmé par le fait que des adverbes temporaux dénotant des moments distincts sont permis avec ces verbes.

(7) a. Jean souhaite lire ce livre.
b. Jean constate avoir lu ce livre.

(8) a. Hier, Jean a souhaité lire ce livre aujourd'hui.
b. Demain, Jean constatera avoir lu ce livre aujourd'hui.

Il semble donc que d'un point de vue sémantique les verbes effectifs jouent essentiellement un rôle de modificateurs face à l'événement dénoté par le prédicat enchâssé et ses arguments.

### 2.2.2. Les propriétés syntaxiques des complétives

Sur le plan syntaxique, les constructions dans lesquelles apparaissent les prédicats effectifs présentent également un certain nombre de propriétés qui les distinguent des constructions infinitives introduites par des prédicats émotifs ou propositionnels.

De façon générale, les compléments infinitifs des prédicats effectifs ne se comportent pas comme des objets de ces verbes en ce sens qu'ils ne peuvent ni être questionnés directement ni être remplacés par un pronom anaphorique, contrairement aux compléments des prédicats émotifs et propositionnels.

Par exemple, les questions introduites par *qu'est-ce que* ne sont acceptables que si elles portent sur des objets nominaux des verbes effectifs. En réponse aux questions données en (9), seules les phrases en (10) sont véritablement acceptables et non celles données en (11). En fait, il est nécessaire d'avoir recours au verbe de rappel *faire* pour questionner le complément infinitif des verbes effectifs, comme le montrent les questions en (12) qui peuvent recevoir comme réponse les phrases en (11).

(9) a. Qu'est-ce que Jean a commencé ?
b. Qu'est-ce que Jean a essayé ?

(10) a. Jean a commencé la lecture du livre.
b. Jean a essayé une nouvelle tactique.

(11) a. Jean a commencé à lire le livre.
b. Jean a essayé de mettre en place une nouvelle tactique.

(12) a. Qu'est-ce que Jean a commencé à faire ?
b. Qu'est-ce que Jean a essayé de faire ?

Les verbes émotifs et propositionnels permettent par contre de questionner directement la complétive qu'ils introduisent, comme le montrent les exemples suivants.

(13) a. Qu'est-ce que Jean souhaite ? Jean souhaite lire le livre.
b. Qu'est-ce que Jean affirme ? Jean affirme être le meilleur.

Par ailleurs, les compléments infinitifs des verbes effectifs ne peuvent être remplacés par un pronom anaphorique, comme l'illustrent les exemples en (14) avec dislocation à droite. La seule référence possible pour un pronom accusatif utilisé avec ces verbes est la référence à un syntagme nominal objet direct, dans la mesure où un tel complément est permis par le verbe. Une fois de plus, le contraste avec des verbes émotifs ou propositionnels, illustré en (16), est révélateur.

(14) a. *Jean l'a commencé, à lire le livre.
b. *Jean l'a essayé, de partir.

(15) a. Jean l'a commencée, la lecture du livre.
b. Jean l'a essayée, la nouvelle tactique.

(16) a. Jean le souhaitait vivement, (de) partir à l'étranger.
b. Jean l'a affirmé maintes fois, être le meilleur.

L'impossibilité de référer au complément infinitif des verbes effectifs au moyen d'un pronom ou de le questionner directement ne peut être attribuée au fait que celui-ci est introduit par une préposition, puisque d'autres occurrences de compléments infinitifs introduits par des prépositions peuvent être remplacées par un pronom ou être questionnées.

(17) a. Jean le lui avait permis, d'aller au cinéma.
b. Qu'est-ce que Jean a permis à Pierre ? Jean lui a permis d'aller au cinéma.

Ces deux tests syntaxiques concordent donc à montrer que les compléments infinitifs des verbes effectifs ne se comportent pas comme ceux des prédicats émotifs ou propositionnels et que, contrairement à ces derniers, ils ne semblent pas entrer dans une relation d'objet face aux verbes matrices qui les introduisent.

Dans des langues comme l'italien et l'espagnol, les compléments infinitifs de certains prédicats effectifs se distinguent également de ceux des autres verbes matrices par leur transparence face à certains processus syntaxiques tels que la montée des clitiques, comme le montrent les exemples suivants tirés de Rizzi (1978).

(18) a. Gianni gli continua a raccontare stupide storie.
'Gianni continue à lui raconter des histoires stupides.'
b. Piero li venne a chiamare alla stazione.
'Piero est venu les chercher à la gare.'

Dans Rochette (1988, 1990), j'ai montré que la classe des verbes qui peuvent ainsi participer à la formation de prédicats complexes syntaxiques est constituée d'un sous-ensemble de la classe des prédicats effectifs. Il s'agit en particulier des verbes aspectuels, modaux et de mouvement. J'ai argumenté que la formation de prédicats complexes est fonction, d'une part, du fait que ces verbes ne peuvent pas constituer des événements autonomes et, d'autre part, du fait que leurs compléments infinitifs se réalisent syntaxiquement comme une projection verbale plutôt que phrastique.

### 2.3. Résumé

Les arguments avancés en faveur d'une catégorie sémantique *action* reposent donc essentiellement sur le fait que les prédicats effectifs ne peuvent sélectionner que des complétives à l'infinitif et que celles-ci ne constituent pas des événements autonomes, ne se comportent pas comme des objets et peuvent même former des prédicats complexes avec le verbe matrice, contrairement aux complétives infinitives des verbes émotifs ou propositionnels.

### 3. La catégorie *action* est-elle vraiment justifiée ?

Il demeure toutefois que sur le plan sémantique l'existence d'une catégorie sémantique *action* semble en fait plutôt douteuse tout comme son opposition éventuelle à celle d'*événement*.

J'aimerais maintenant me pencher sur le statut sémantique de la catégorie *action* et par la suite réexaminer les propriétés des prédicats effectifs et de leurs complétives afin de déterminer si ces propriétés ne relèveraient pas d'autres facteurs.

#### 3.1. Le statut de la catégorie sémantique *action*

Une première objection que l'on pourrait être tenté d'apporter à l'existence d'une catégorie *action* est que celle-ci devrait s'opposer à celle d'*état* plutôt qu'à celle d'*événement*. Mais cette objection ne constitue pas un véritable contre-argument puisque l'opposition entre les différentes catégories sémantiques ne doit pas nécessairement être de type binaire. En effet, l'opposition entre les catégories *événement* et *proposition* ne semble pas être de type binaire mais plutôt de type «hiérarchique» en ce sens que les *propositions* renferment elles-mêmes des *événements*. On peut dès lors envisager que l'opposition entre les *actions* et les *événements* est du même ordre et que les *actions* constituent en fait une sous-classe d'*événements*.

Un argument additionnel avancé à l'origine en faveur de l'existence de la catégorie *action* trouve sa source dans le fait que plusieurs verbes effectifs ne tolèrent pas facilement la présence de verbes statifs dans leurs complétives (Gross 1968, 1975, Lamiroy 1987).

(19) *Jean commence à avoir une maison.

Il y aurait donc lieu de caractériser leur sélection sémantique comme plus restreinte qu'un *événement*, et comme non stative, de là le recours à la catégorie *action*.

Mais cet argument se heurte au fait que ce ne sont pas tous les verbes effectifs qui présentent ce type de restrictions de sélection. Pour ceux qui ne présentent pas de restrictions de sélection (p.ex. les modaux *pouvoir* et *devoir*), il est difficile d'arguer que ces verbes

sélectionnent une catégorie autre que celle d'*événement* (Boivin 1994, Boivin et Rochette 1994).

(20) Jean peut avoir une maison.

En fait, les restrictions de sélection observées dans le cas de verbes aspectuels comme *commencer* se révèlent ne pas être des restrictions opposant *actions* et *états*. Dans Rochette (1993), j'ai démontré que les verbes aspectuels sélectionnent des activités ou des accomplissements et que de façon générale ils ne permettent pas les achèvements ou les états. J'ai proposé de caractériser les verbes aspectuels comme sélectionnant la catégorie sémantique *processus*, qui recouvre les activités et les accomplissements.

Par ailleurs, les classifications aspectuelles les plus couramment utilisées, basées principalement sur les travaux de Vendler (1967), ne distinguent généralement pas une catégorie *action*. Tout ceci soulève donc de façon cruciale la pertinence de maintenir l'existence de cette catégorie en tant que primitif sémantique.

### 3.2. Les propriétés des verbes effectifs

Comme je l'ai souligné en 2.2 ci-dessus, l'impossibilité d'utiliser des adverbiaux temporels distincts démontre essentiellement que les verbes effectifs ne peuvent constituer un événement autonome par rapport à celui dénoté par leur complétive. Cet argument peut donc être vu comme un argument qui démontre que les prédicats effectifs entretiennent avec leurs complétives un rapport tel que celles-ci ne peuvent être dissociées du verbe matrice plutôt que comme un argument en faveur d'une catégorie sémantique distincte.

Dans la même veine, nous pouvons également réinterpréter les propriétés syntaxiques mises à jour par l'emploi de tests syntaxiques comme l'interrogation, la pronominalisation ou la montée des clitiques comme des propriétés qui nous renseignent surtout sur le rôle joué par le prédicat matrice plutôt que sur la nature sémantique de la complétive infinitive. Par exemple, le test de l'interrogation nous montre que le prédicat matrice n'a pas de statut autonome par rapport à la complétive qu'il introduit. Il semble en fait constituer un «prédicat complexe» avec la complétive puisqu'il est nécessaire d'avoir recours à un verbe factice comme *faire* pour questionner l'événement dénoté par la complétive.

En fait, ce qui ressort le plus clairement des différentes propriétés examinées, c'est que les prédicats effectifs semblent jouer un rôle de «semi-auxiliaires» ou «modificateurs» face à la «complétive» qu'ils introduisent. Ceci est nettement le cas lorsqu'il s'agit des verbes aspectuels et modaux, qui apportent l'aspect ou la modalité à l'événement dénoté par le prédicat enchâssé. Mais ceci est vrai également de verbes tels que *oser* qui apporte une modification plutôt de type adverbial.

En ce sens, les verbes effectifs se rapprocheraient donc à la fois des verbes auxiliaires et des adverbes. Ils partageraient avec ces deux catégories la propriété de ne pas être les «prédicateurs primaires» de la phrase. Tout comme les auxiliaires, ils apparaissent toutefois comme l'élément verbal porteur des marques de flexion et tout comme les adverbes, ils contribuent une forme de prédication secondaire à l'événement dénoté par le prédicat infinitif et ses arguments.

Ce qui resterait à élucider maintenant c'est la forme sous laquelle on doit encoder la sélection par les verbes effectifs. En d'autres termes, que doit-on spécifier exactement dans les entrées lexicales des verbes pour garantir qu'il sera possible de distinguer un verbe effectif comme *oser* d'un verbe émotif comme *désirer* si tous deux sélectionnent la même catégorie *événement*. Car, après tout, la postulation de la catégorie sémantique *action* avait le mérite de pouvoir rendre compte des différences de comportement de ces verbes.

## 4. Conclusion

Les différents arguments présentés précédemment montrent qu'il n'est pas clair que les prédicats effectifs sélectionnent une catégorie sémantique unique et distincte de celle sélectionnée par les prédicats émotifs. Dans certains cas, notamment dans celui des verbes aspectuels, il semble que la catégorie sélectionnée corresponde à une catégorie plus «restreinte» que celle d'*événement*, notamment celle de *processus*. Mais dans d'autres cas, comme dans celui des verbes modaux qui ne présentent pas de telles restrictions de sélection, la catégorie *événement* semblerait appropriée pour caractériser le type sémantique de la complétive.

Les arguments en faveur de la catégorie *action*, basés sur les propriétés sémantiques et syntaxiques des structures dans lesquelles apparaissent les prédicats effectifs, peuvent être réinterprétés comme des arguments qui soulignent essentiellement le rôle de «semi-auxiliaires» de ces verbes. Il serait ainsi possible de caractériser la sélection des verbes effectifs en termes de la catégorie *événement* et de faire porter le poids des différences de propriétés sémantiques et syntaxiques aux verbes effectifs eux-mêmes plutôt que de postuler l'existence d'une catégorie sémantique *action*.

## Bibliographie

Boivin, Marie Claude. 1994. *Les interprétations radicale et épistémique des verbes* devoir *et* pouvoir. Mémoire de maîtrise, UQAM.

Boivin, Marie Claude, et Anne Rochette. 1994. "*Pouvoir* et *devoir* dans les complétives infinitives." In P. Koskinen (réd.) *Actes du Congrès annuel de l'Association canadienne de linguistique* 1994. Toronto Working Papers in Linguistics, University of Toronto, pp. 51-62.

Grimshaw, Jane. 1979. "Complement Selection and the Lexicon". *Linguistic Inquiry* 10: 279-426.

Grimshaw, Jane. 1981. "Form, Function and the Language Acquisition Device". In C.L. Baker & J. McCarthy (réd.) *The Logical Problem of Language Acquisition*. Cambridge, Mass.: MIT Press.

Gross, Maurice. 1968. *Grammaire transformationnelle du français: syntaxe du verbe*. Paris: Larousse.

Gross, Maurice. 1975. *Méthodes en syntaxe*. Paris: Hermann.

Lamiroy, Béatrice. 1987. "The complementation of aspectual verbs in French". *Language* 63: 278-298.

Long, Mark. 1974. *Semantic Verb Classes and their Role in French Predicate Complementation*. Thèse de Ph.D., Indiana University, Bloomington.

Melvold, Janice. 1986. "Factivity and Definiteness". Ms., MIT.

Pesetsky, David. 1982. *Paths and Categories*. Thèse de Ph.D., MIT.

Rizzi, Luigi. 1978. "A Restructuring Rule in Italian Syntax". In S. J. Keyser (réd.) *Recent Transformational Studies in European Languages*. Cambridge, Mass.: MIT Press, pp. 113-158.

Rochette, Anne. 1988. *Semantic and Syntactic Aspects of Romance Sentential Complementation*. Thèse de Ph.D., MIT.

Rochette, Anne. 1990. "On the Restructuring Classes of Verbs in Romance". In A.-M. Di Sciullo & A. Rochette (réd.) *Binding in Romance. Essays in Honour of Judith McA'Nulty*. Ottawa: Publication spéciale de l'Association canadienne de linguistique, pp. 96-128.

Rochette, Anne. 1991. "La structure d'arguments et les propriétés distributionnelles des adverbes". *Revue québécoise de linguistique* 20.1: 55-78.

Rochette, Anne. 1993. "À propos des restrictions de sélection de type aspectuel dans les complétives infinitives du français". *Langue française* 100: 67-82.

Vendler, Zeno. 1967. *Linguistics and Philosophy*. Ithaca: Cornell University Press.

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 89-103.

# Les connecteurs pragmatiques comme indices du développement du langage chez les enfants de 5 à 11 ans

*Madeleine Saint-Pierre*
Université du Québec à Montréal

## 1. Objectifs de l'étude

Dans cette étude j'ai voulu explorer des aspects pragmatiques et développementaux du lexique en décrivant l'emploi des connecteurs dans le discours d'enfants de 5 à 11 ans. Les connecteurs sont analysés d'un point de vue sémantico-pragmatique comme marqueurs de cohésion argumentative et de structuration du dialogue et d'autre part, d'un point de vue développemental comme indices du développement des capacités de communication des enfants de cet âge. Les données extraites d'un corpus de conversations dyadiques entre enfants francophones de Montréal permettront de vérifier si les connecteurs font partie des traits du langage qui se développent après l'âge de 5 ans.

## 2. Cadre théorique et problématique

Ce travail s'insère dans une recherche portant sur le développement des compétences pragmatiques du langage chez les enfants dont le principal objectif est de décrire les aspects du langage qui se développent une fois le système grammatical acquis, c'est-à-dire après l'âge de 5 ans; cf. Saint-Pierre (1989; 1995) et Saint-Pierre et Langevin (1990). En ce qui concerne les aspects développementaux, notre étude prend appui sur un ensemble de travaux relevant des domaines suivants: la sociocognition, dans la perspective des travaux de Lefebvre-Pinard *et al* (1982) et Lefebvre-Pinard (1984; 1985) qui font état de changements importants sur le plan du développement sociocognitif des enfants entre l'âge de 5 et 10 ans; la psycholinguistique génétique avec les travaux de Baudichon (1982); Bernicot (1992) et la pragmatique développementale tel que dans les études d'Ochs et Schieffelin (1979) et de Shatz (1978) principalement.

Par ailleurs, les descriptions et catégorisations des connecteurs s'inspirent des travaux de l'École de Genève et principalement de ceux de Roulet, Auchlin, Moeschler (1985); Moeschler (1989); Lusher (1994) et Moeschler et Reboul (1994). Selon ces études, les connecteurs sont définis dans leur dimension pragmatique en tant manifestation des liens de cohésion entre les actes de langage ou les interventions dans une conversation. Ils sont

distincts des marqueurs de structuration de la conversation (MSC) définis dans Roulet et al. (1985) comme des signaux de structuration et selon Saint-Pierre et Vadnais (1992), comme des marques de ponctuation de l'oral spontané ayant surtout une fonction de marqueur d'interaction verbale. Les connecteurs pragmatiques ont la propriété d'une part de connecter des actes de langage et d'autre part, de réaliser des actes d'argumentation, c'est-à-dire des actes obligeant l'interlocuteur à interpréter les énoncés comme autant d'arguments pour certaines conclusions (généralement implicites) visées par le locuteur; cf. Moeschler (1989); Moeschler et Reboul (1994); Lusher (1994).

Les connecteurs pragmatiques qui ont une fonction argumentative sont regroupés en 4 types selon Roulet et al. (1985: 112); soit:

1° les connecteurs **argumentatifs** qui introduisent les arguments et les relient à l'acte directeur de la séquence: "*parce que, puisque, en effet* " qui prennent un seul argument, ainsi que "*même, d'ailleurs* " qui en prennent deux.

2° les connecteurs **contre-argumentatifs** qui introduisent un contre-argument en regard de l'acte directeur: "*bien que, malgré (que), même si, quand même, cependant, toutefois...*"

3° les connecteurs **consécutifs** qui introduisent une relation de conséquence entre l'argument et l'acte directeur tels que: "*donc, par conséquent, aussi, ainsi....*"

4° les connecteurs **réévaluatifs** qui servent à résumer, récapituler qui peuvent être co-orientés en regard de la conclusion tels que: "*bref, en somme, au fond, décidément* " ou anti-orientés en regard de la conclusion "*en fin de compte, finalement, de toute façon...*".

Les connecteurs à fonction argumentative de notre corpus ont été classés dans l'un de ces 4 sous-types alors que les connecteurs à valeur non-argumentative ont été classés sur une base sémantique selon la classification proposée par Lundquist (1994:51).

La question centrale de ce travail est donc de savoir si la production de telles marques de connexion qui explicitent la cohérence interne du discours augmente et se diversifie avec l'âge. Cette description complète les résultats de McTear (1985) et notre propre étude sur le développement de la cohérence conversationnelle; cf. Saint-Pierre et Langevin (1990).

Notre analyse propose donc de décrire les aspects sémantiques et pragmatiques des connecteurs en considérant d'abord

la distribution des connecteurs produits à chaque âge et la diversité des emplois d'un même connecteur d'autre part.

## 3. Méthode

Le corpus Feider-Saint-Pierre de conversations dyadiques semi-structurées entre enfants regroupés par affinité naturelle et selon leur âge a servi de base à nos analyses. Les locuteurs franco-montréalais sont issus d'un milieu socio-économique moyen et fréquentent la maternelle, troisième et cinquième année de l'école de quartier. L'expérimentatrice propose à chaque dyade d'enfants de faire un jeu de rôle sur le thème "une journée d'un ou d'une vétérinaire et de son assistant ou assistante dans leur clinique". Le corpus vidéofilmé a été transcrit et informatisé et un code de transcription a été élaboré pour consigner les données verbales propres à l'oral (linguistiques, paralinguistiques et discursives) et les comportements non-verbaux pertinents pour l'analyse des conversations. L'unité de segmentation du discours dans la transcription est l'acte de langage. L'acte de langage, rappelons-le, est l'acte que l'on accomplit **en** disant quelque chose. Il est défini dans la théorie des actes de langage sur la base d'une seule proposition ou contenu propositionnel. Cette théorie est résumée dans Moeschler et Reboul (1994: chap.1)

Pour analyser l'emploi des connecteurs, nous avons extrait du corpus Feider-Saint-Pierre l'échantillon suivant, soit: 20 enfants dont l'âge moyen est de 5;8 ans répartis en 3 dyades de filles et 5 dyades de garçons; 20 enfants de 8;8 ans dont 4 dyades de filles, 5 de garçons et 1 dyade mixte puis, 20 enfants de 10;8 ans d'âge moyen répartis en 4 dyades de filles, 5 dyades de garçons et 1 dyade mixte. Nous avons extrait 75 pages de dialogue de la transcription informatisée pour chaque groupe d'âge dans l'épisode du jeu de rôles qui consiste à soigner les animaux afin d'obtenir des données comparables qualitativement et quantitativement pour chaque groupe d'âge.

Les étapes de l'analyse sont les suivantes: 1° le dépouillement des marques de connexion telles que définies plus haut appartenant à diverses catégories grammaticales: conjonctions de coordination et de subordination, adverbes et locutions diverses. 2° La classification des connecteurs en distinguant les 4 valeurs argumentatives présentées plus haut et les valeurs non-argumentatives suivantes: temporalité, comparaison, énumération, addition et alternance.

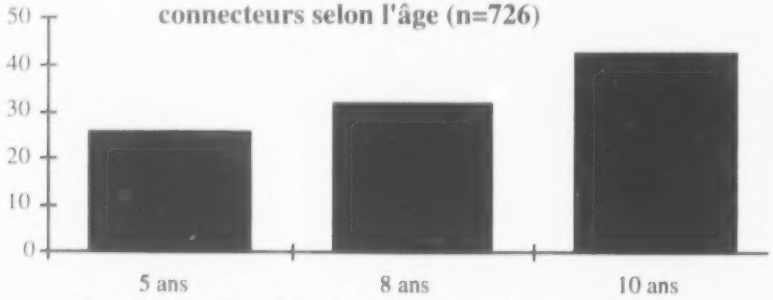
Les figures présentées ci-après illustrent donc la distribution des différents types de connecteurs selon des variables telles que

l'âge et le sexe des locuteurs afin de répondre à la question relative au développement de la compétence à communiquer des enfants posée initialement .

## 4. Résultats et discussion

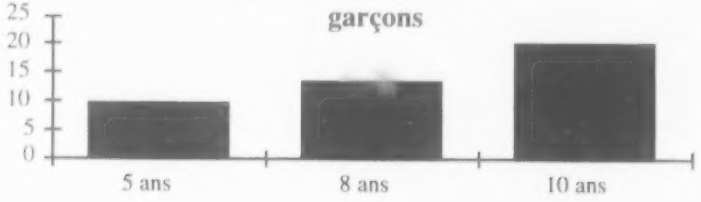
Nous avons repéré 726 connecteurs dans ce corpus. Toutes catégories de connecteurs confondues, la figure 1 indique une nette augmentation du nombre de connecteurs selon l'âge passant de 25,6% à 5 ans, à 32% à 8 ans et 42, 4% à 10 ans.

**Fig.1 Répartition du nombre moyen de connecteurs selon l'âge (n=726)**




Voyons la répartition du nombre des connecteurs selon le sexe pour chaque groupe d'âge dans les figures 2 et 3 qui suivent.

**Fig.2 Répartition du nombre moyen de connecteurs chez les garçons**

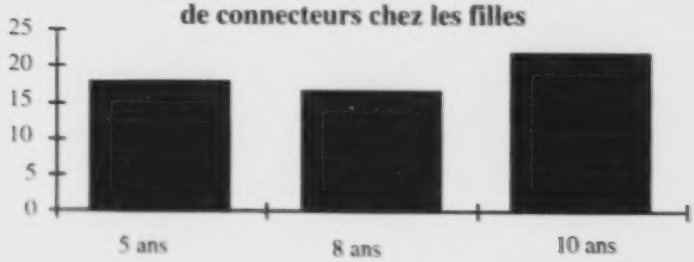



La figure 2 montre que dans le groupe des garçons de 5 ans, il y a 9,8% des connecteurs soit en moyenne 15 connecteurs par dyade; dans le groupe de 8 ans, une moyenne de 23 connecteurs par dyade soit 13,6 % et dans le groupe de 10 ans, une moyenne de 31

connecteurs par dyade soit 20,4 % de l'ensemble. L'augmentation du nombre de connecteurs est constante d'un groupe d'âge à l'autre.

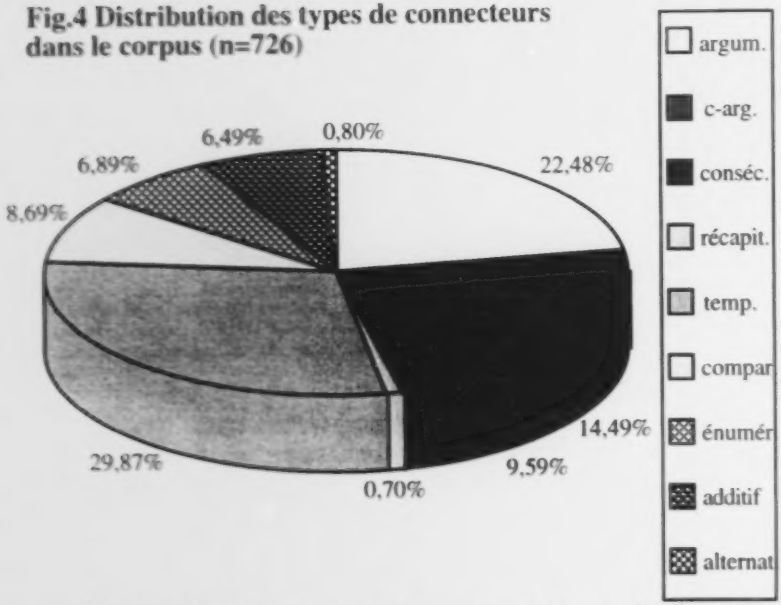
La figure 3 montre la répartition du nombre de connecteurs dans le groupe des filles selon leur âge. Les connecteurs sont plus nombreux dans les conversations des filles que dans celles des garçons et ce, sutout à l'âge de 5 ans. En effet, on observe une moyenne de 27 connecteurs à 5 ans, soit 17,9%, une moyenne de 25 connecteurs, soit 16,4% à 8 ans et une moyenne de 31 connecteurs par dyade soit 22% à 10 ans. En termes de fréquence, la différence entre les filles et les garçons s'amenuise avec l'âge. À 10 ans, les filles et les garçons produisent le même nombre de connecteurs en moyenne.

**Fig.3 Répartition du nombre moyen de connecteurs chez les filles**




La figure 4 montre la distribution de tous les types de connecteurs recueillis dans le corpus ainsi que leurs proportions relatives.

**Fig.4 Distribution des types de connecteurs dans le corpus (n=726)**

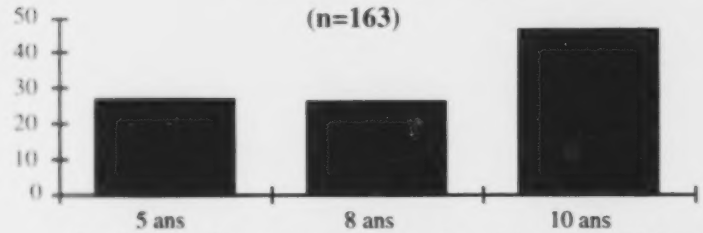



Nous observons que les connecteurs à fonction **argumentative** représentent 47% soit 343 des 726 connecteurs du corpus dont 22,5% sont des argumentatifs, 14,5% des contre-argumentatifs, 9,6% des consécutifs et 0,7% des récapitulatifs. Les connecteurs **non-argumentatifs** se répartissent de la façon suivante: 29,9% temporels, 8,7% comparatifs, 6,9% énumératifs, 6,5% additifs et 0,8% alternatifs pour un total de 52,7% de l'ensemble.

La série de figures suivantes permettra de voir si la répartition des connecteurs croît selon l'âge des sujets tel que l'avons supposé initialement. Pour ce faire, nous avons distingué les connecteurs à fonction argumentative des connecteurs à fonction non-argumentative. Les figures 5 à 8 illustrent la répartition des 343 connecteurs à fonction **argumentative** selon l'âge. La figure 5 montre une augmentation du nombre de connecteurs proprement argumentatifs entre l'âge de 8 ans et 10 ans seulement. Une analyse plus détaillée des argumentatifs montre que dès l'âge de 5 ans quatre types sont déjà employés.pour exprimer les dimensions suivantes: justification ou causalité, explication, condition, hypothèse et finalité

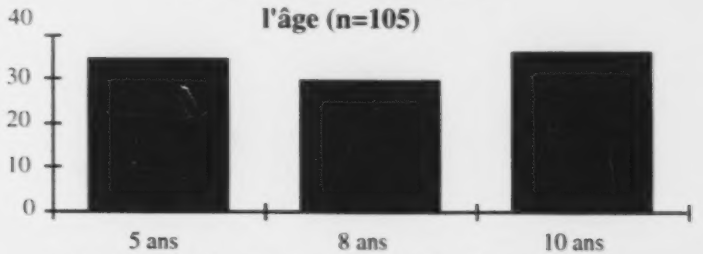
ou but. La distribution des fréquences relatives à ces emplois apparaissent au tableau 1 de l'annexe.

**Fig.5 Distribution des connecteurs argumentatifs selon l'âge (n=163)**




La figure 6 montre que la répartition des 163 contre-argumentatifs n'augmente pas selon l'âge et même qu'ils sont utilisés un peu plus souvent chez les enfants de 5 ans que chez les enfants de 8 ans. Les données présentées dans le tableau 2 en annexe montrent que c'est *mais* qui est de loin le plus souvent utilisé pour contre-argumenter et ce, quel que soit l'âge des sujets. En effet, dès l'âge de 5 ans, ce type de connecteur sert à trois fonctions: l'opposition, la restriction et la concession. Les données du tableau 2 montrent également qu'il n'y a pas de diversification des connecteurs contre-argumentatifs selon l'âge. De plus, le connecteur *mais* sert à introduire un acte directif, question ou une demande d'action, agissant ainsi comme marqueur de prise de parole.

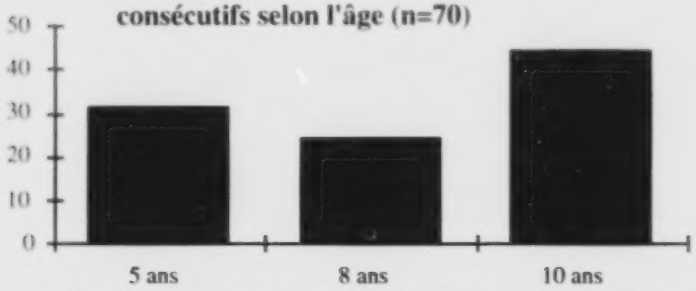
**Fig.6 Distribution de connecteurs contre-argumentatifs selon l'âge (n=105)**




La figure 7 montre la distribution des 70 connecteurs consécutifs dans le discours des enfants. On peut observer une

augmentation de ces connecteurs entre l'âge de 8 et 10 ans. Le tableau 3 de l'annexe montre que la diversité des connecteurs de conséquence est aussi plus grande à 10 ans que dans les autres groupes d'âge. Il en est de même pour les conclusifs. On remarque en effet que seuls les enfants de 10 ans ont utilisé *donc* ou *d'abord* pour conclure leur intervention.

**Fig.7 Distribution de connecteurs consécutifs selon l'âge (n=70)**




Le quatrième type de connecteur à fonction argumentative est le récapitulatif. Ce connecteur est très rare dans notre corpus car seulement 5 de ces connecteurs ont été observés et tous ont été produits par les enfants de 10 ans. Il s'agit du connecteur *en tout cas*.

La répartition des types de connecteurs à fonction argumentative dans notre corpus montre qu'argumenter en expliquant, justifiant, posant des conditions, des hypothèses ou en exprimant des buts à l'aide de connecteurs divers est une compétence déjà bien installée dans le discours des enfants dès l'âge de 5 ans. Toutefois, les enfants de 10 ans emploient un plus grand nombre de ces connecteurs que leurs cadets. Pour ce qui est de la capacité à contre-argumenter même si le nombre de connecteurs contre-argumentatifs est un peu moins important dans le corpus, on remarque que les enfants de 5 ans emploient autant de connecteurs d'opposition et de restriction que leurs aînés et moins de connecteurs de concession que ces derniers. Ceci démontre une capacité à l'argumentation de type dialectique déjà bien en place à cet âge. Quant au recours aux connecteurs consécutifs, on note que leur diversité augmente avec l'âge et ce, autant pour exprimer un lien de conséquence qu'une conclusion, tel qu'illustré dans le tableau 3 en annexe. La rareté des connecteurs récapitulatifs peut s'expliquer, puisqu'ils sont cognitivement plus difficiles à produire à cause du fait

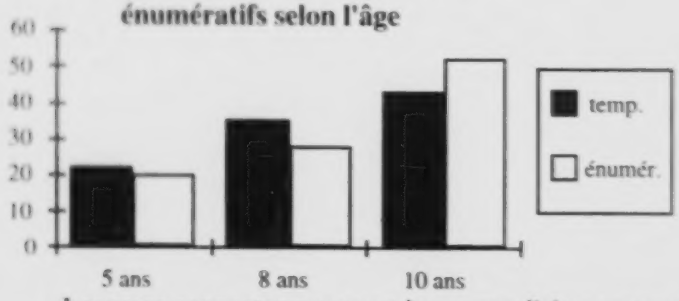
qu'ils introduisent un argument dont l'orientation ne va pas dans le même sens que la conclusion, comme c'est le cas de *en tout cas*.

La deuxième partie de nos analyses présente la distribution des divers connecteurs à fonction non-argumentative. À l'instar des connecteurs à fonction argumentative, voyons s'ils sont fonction de l'âge des sujets.

Les connecteurs **non-argumentatifs** du corpus sont par ordre de fréquence: les temporels (n=217), les comparatifs (n=63), les énumératifs (n=50), les additifs (n=47) et les alternatifs (n=6).

La figure 8 montre que l'emploi des connecteurs temporels et des énumératifs augmente nettement selon l'âge des sujets. Le développement est marqué par deux dimensions: l'augmentation du nombre des connecteurs dans les deux types et également, la diversification des connecteurs de chaque type . Les connecteurs temporels produits uniquemqnt par les enfants de 10 ans sont: *pendant que*, *min que*, *pendant ce temps*, *pour l'instant*, *en même temps que*, et *pour quand que*. Plusieurs de ces connecteurs introduisent une subordonnée circonstancielle qui, sur le plan syntaxique, correspond à une construction complexe plus difficile à produire. Le degré de difficulté est d'ordre syntaxique et sémantique.

**Fig.8 Connecteurs temporels et énumératifs selon l'âge**

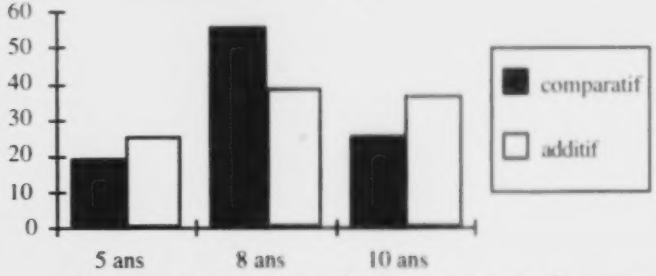



Les rares connecteurs servant à marquer l'alternance entre les énoncés tels que *ou* et *ou bien* (n=6) ont été utilisés en exclusivité par les enfants de 10 ans. On ne peut éviter le lien à établir avec la présence exclusive des 5 connecteurs récapitulatifs du corpus dans le langage des enfants de 10 ans. Cette rareté serait-elle synonyme de difficulté? Il serait prématuré d'y voir un indice développemental à

cause du peu de données existantes mais il y aurait lieu de vérifier cette tendance dans un corpus plus large.

Pour terminer, voyons la distribution des fréquences des marqueurs de comparaison et d'addition dans la figure 9 ci-après.

**Fig.9 Distribution des comparatifs et des additifs selon l'âge**




Nous observons que les connecteurs de comparaison et d'addition sont employés à 5 ans dans les mêmes proportions relatives qu'à 10 ans avec une moins grande fréquence cependant. De façon plutôt surprenante, ce sont les enfants de 8 ans qui ont le plus souvent recours à ces types de connecteurs. Il faut se rappeler que cette répartition rappelle celles des connecteurs contre-argumentatifs et consécutifs présentées ci-haut. On observe que les moyens pour exprimer des comparaisons sont déjà bien en place dès l'âge de 5 ans. Les marqueurs tels que: *comme, même que, pire que, mieux que, plus que, moins que...* sont en effet employés couramment à cet âge.

En ce qui concerne les connecteurs additifs, les formes les plus fréquemment observées sont: *et, puis, aussi,* employés chez les enfants dès l'âge de 5 ans mais moins fréquemment que chez les enfants plus âgés. La courbe de fréquence observée dans ce cas est assez conforme à l'ensemble des autres résultats de notre recherche qui montrent une progression nette de certains traits entre l'âge de 5 et de 8 ans pour se stabiliser entre l'âge de 8 et de 10 ans.

## 5. Conclusion

Les résultats de cette étude montrent qu'un ensemble important de connecteurs est déjà maîtrisé à l'âge de 5 ans et ce, davantage chez les filles que chez les garçons de cet âge. Proportionnellement, les enfants de 5 ans ont utilisé moins de

connecteurs à fonction **non-argumentative** (20%) que de connecteurs à fonction **argumentative** (environ 30%). Selon les types de connecteurs cette proportion est soit égale, soit supérieure au pourcentage des connecteurs produits par les enfants de 8 ans. En effet, si l'on regarde les fréquences d'utilisation des 4 types de connecteurs à fonction argumentative, la progression est nette entre l'âge de 8 et de 10 ans à une exception près: les contre-argumentatifs. Dans ce dernier cas, les proportions se situent autour de 30 à 35% pour les trois groupes d'âge et les connecteurs employés sont sensiblement les mêmes dans les trois groupes.

Si le nombre de connecteurs augmente avec l'âge, il y a aussi une plus grande diversité des connecteurs servant une même fonction selon l'âge. C'est là que se trouve le deuxième aspect du comportement linguistique en évolution. On observe en effet, une plus grande variété de connecteurs proprement argumentatifs employés pour marquer une relation de causalité, de condition, de finalité ou d'explication chez les enfants de 10 ans. De même, les connecteurs consécutifs tant pour exprimer la conséquence que la conclusion d'une argumentation sont plus variés à 10 ans que dans les deux autres groupes d'âge. (Tableaux 1 et 3 de l'annexe).

Par ailleurs, les connecteurs à fonction **non argumentative** sont moins souvent utilisés par les enfants de 5 ans (20%) que par les enfants de 8 (30% à 40%) ou de 10 ans (30 à 50 %). Ceci est vrai pour les connecteurs temporels, énumératifs, comparatifs et additifs. Non seulement le nombre de connecteurs augmente avec l'âge mais on constate également une plus grande variété de connecteurs temporels dans le langage des enfants plus âgés au nombre de 12 à 10 ans, de 8 à 8 ans et de 5 à 5 ans.

À l'instar des résultats des études de Shatz (1978), Bates et Mac Whinney (1982), Beaudichon (1982) et Dore (1985) et en conformité avec nos propres travaux sur le développement des actes de langage, des formes de politesse ou celui de la cohérence conversationnelle, nous constatons que les enfants de 5 ans possèdent un niveau de compétence comparable à maints égards à celui de leurs aînés de 8 et 10 ans. Une explication de ce niveau de compétence pourrait se trouver dans le fait que la situation de jeux de rôles suscite l'emploi de structures et de formes linguistiques variées et élaborées de la part de l'enfant telles que l'adulte qu'il imite les produirait. Cette situation permet aux enfants de parler autrement qu'ils ne le feraient avec leurs pairs d'une part et, d'autre part, elle ne les contraint pas comme le font les situations expérimentales ou

didactiques qui sont perçues par plusieurs chercheurs comme limitatives à bien des égards; cf. Lefebvre-Pinard (1984).

REFERENCES BIBLIOGRAPHIQUES

Bates, Elizabeth & Brian Mac Whinney. 1982.*Language Acquisition: the state of the art*. Cambridge: Cambridge University Press.

Beaudichon, Janine. 1982. *La communication sociale chez l'enfant*. Paris: Presses Universitaires de France.

Bernicot, Josie. 1992. *Les actes de langage chez l'enfant*. Paris: Presses Universitaires de France.

Dore, John. 1985. "The development of conversational competence". In R.L. Schiefelbush. (réd.). *Communication competence: assessment and intervention*. San Diego: College Hill Press, pp. 3-57.

Lefebvre-Pinard, Monique, Thérèse Bouffard-Bouchard & Helga Feider. 1982. "Social cognition and verbal requests among preschool children". *The Journal of Psychology* 110: 133-143.

Lefebvre-Pinard, Monique. 1984. "Les relations entre la cognition sociale et les interactions sociales chez le jeune enfant: un examen critique". In Deleau, Michel (réd.). *Langage et Communication à l'âge préscolaire*. Rennes: Presses Universitaires de Rennes 2, pp. 137-150.

Lefebvre-Pinard, Monique. 1985. "La régulation de la communication de l'enfance à l'âge adulte". In Noizet, Georges et al. (réd.). *La Communication*. Paris: Presses de l'Université de France, pp. 107-132.

Lundquist, Lita. 1994. *La cohérence textuelle*. Frederiksberg C.: Samfundslitteratur.

Luscher, Jean-Marc. 1994. "Les marques de connexion: des guides pour l'interprétation". In J. Moeschler et al. *Langage et pertinence*. Nancy: Presses Universitaires de Nancy. Collection Processus Discursif, pp. 175-228.

McTear, Michael F. 1985. *Children's conversation*. Oxford: Blackwell.

Moeschler, Jacques. 1989. *Modélisation du dialogue, représentation de l'inférence argumentative*. Paris: Hermes.

Moeschler, Jacques & Anne Reboul. 1994. *Dictionnaire encyclopédique de pragmatique*. Paris: Seuil.

Ochs, Elinor & Bambi B. Schieffelin. 1979. *Developmental Pragmatics*. New York, San Francisco, London: Academic Press.

Roulet, Eddy, Antoine Auchlin & Jacques Moeschler. 1985. *L'articulation du discours en français contemporain*. Berne, Frankfort-s. Main, New York: Peter Lang.

Saint-Pierre, Madeleine. 1989. "Les actes de langage comme indicateurs du statut social entre enfants en situation de jeu de rôle". *Revue québécoise de linguistique* 18, 1: 129-146.

Saint-Pierre, Madeleine. 1991. "Illocutoire et modalisation: les marqueurs d'intensité en français". *Revue québécoise de linguistique* 20, 2: 223-235.

Saint-Pierre, Madeleine. 1995."Quand dire c'est faire agir: aspects pragmatiques du développement du langage chez les enfants de 5 à 11 ans". Communication présentée au 26e colloque annuel de l'ACLA.

Saint-Pierre, Madeleine & Lysanne Langevin. 1990. "Le courant de la conversation: observations sur le développement de la cohérence dans la conversation d'enfants de 5 et 11 ans". *Revue québécoise de linguistique théorique et appliquée* 9,3:105-128.

Saint-Pierre, Madeleine & Marguerite Vadnais. 1992. "Du modalisateur au marqueur de ponctuation des actions: le cas de *bon*. *Revue québécoise de linguistique* 22,1: 241-255.

Shatz, Marilyn. 1978. "On the Development of Communicative Understandings: An Early Strategy for Interpreting and Responding to Messages". *Cognitive Psychology* 10, 271-301.

ANNEXE

**Tableau 1**
**Connecteurs argumentatifs selon l'âge** (n=163)

| **cause, justification** | **5 ans** | **8 ans** | **10 ans** |
|---|---|---|---|
| parce que | 5 | 5 | 11 |
| à cause que | 0 | 0 | 1 |
| à cause de | 0 | 2 | 3 |
| vu que | 0 | 0 | 3 |
| | | | |
| si (= puisque) | 1 | 0 | 0 |
| tant qu'à (= puisque) | 0 | 0 | 1 |
| | | | |
| | | | |
| **explication** | | | |
| parce que | 13 | 8 | 17 |
| c'est que | 7 | 0 | 3 |
| pour ça que | 0 | 0 | 1 |
| | | | |
| **explication-condition** | | | |
| sinon, sans ça | 2 | 1 | 1 |
| | | | |
| **condition-hypothèse** | | | |
| si, au cas ou | 3 | 12 | 5 |
| peut-être que | 0 | 0 | 1 |
| | | | |
| **finalité-but** | | | |
| pour que | 12 | 15 | 26 |
| pour pas que | 0 | 0 | 3 |

**Tableau 2**
**Connecteurs contre-argumentatifs selon l'âge** (n=105)

| **opposition, adversatif** | **5 ans** | **8 ans** | **10 ans** |
|---|---|---|---|
| mais | 8 | 8 | 9 |
| et (= par contre) | 3 | 0 | 0 |
| pas pour...pour | 0 | 0 | 1 |
| puis (=mais) | 0 | 0 | 2 |
| | | | |
| **restriction** | | | |
| mais | 9 | 2 | 6 |
| par exemple (= toutefois) | 0 | 6 | 4 |
| | | | |
| **concession** | | | |
| mais | 1 | 7 | 5 |
| quand même | 1 | 0 | 4 |
| même si | 1 | 0 | 1 |
| | | | |
| *mais* + question<br>*mais* + demande d'agir | 13 | 8 | 6 |
| | | | |

**Tableau 3**
**Connecteurs consécutifs selon l'âge** (n=70)

| **conséquence** | **5 ans** | **8 ans** | **10 ans** |
|---|---|---|---|
| puis | 13 | 7 | 25 |
| ça fait que | 6 | 2 | 0 |
| comme ça | 0 | 0 | 1 |
| | | | |
| **conclusion** | | | |
| ça fait que | 3 | 3 | 2 |
| alors (= donc) | 1 | 5 | 0 |
| donc | 0 | 0 | 2 |
| d'abord | 0 | 0 | 1 |

# Section 3

## *Aspects sociolinguistiques du lexique*

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 107-119.

# Quels mots les étudiants doivent-ils connaître à leur entrée à l'université?

*Annie Côté et Annette Paquot*
Université Laval

## Introduction

Tout le monde reconnaîtra qu'il est inquiétant que les étudiants qui entrent à l'université ignorent le sens des mots déduction et hypothèse et qu'il est fâcheux pour un journaliste, même étudiant, de confondre pourfendre et défendre. Nous conviendrons sans doute aussi qu'on peut ne pas connaître le sens des mots pentacosiomédime et hyperspénisme et être néanmoins un intellectuel de haut niveau.

Mais peut-on passer de cette appréciation intuitive à une évaluation fondée sur des bases lexicologiques et scolaires solides puisque –on le sait– la compétence lexicale est toujours très partielle et que, reflétant l'expérience de chacun, elle est aussi éminemment variable d'un individu à l'autre?

Officiellement, il semble que oui puisque les tests de français qui ont été administrés depuis quelques années aux étudiants à la fin de leurs études collégiales ont tous compris une évaluation du vocabulaire, que cet aspect de la compétence linguistique a toujours compté pour un pourcentage non négligeable des points et que le système est tel que les résultats obtenus par les étudiants sont déterminants pour la suite de leur parcours scolaire: un échec peut les écarter définitivement des études et une réussite les conforter dans le sentiment qu'ils n'ont aucun problème de cet ordre.

Le test imposé actuellement aux étudiants québécois a été conçu par le ministère de l'Éducation (M.E.Q.). Il consiste en une rédaction dont les critères de correction indiquent que le vocabulaire des étudiants doit être « précis et varié » (M.E.Q. 1992: 32) et pour laquelle ils disposent d'un dictionnaire. Ce test a été précédé par des examens « maison » préparés par les universités se présentant sous la

forme d'examens à correction objective, composés pour l'essentiel de questions à choix multiples dont un certain nombre portaient sur le lexique.

Notre objectif général était de caractériser les aspects de la compétence lexicale qui est l'objet de l'évaluation de tous ces tests, mais le sujet du présent article est plus modeste: nous nous limiterons ici à un seul de ces tests et nous ne considérerons qu'un des aspects de cette compétence, celui qui concerne les mots que l'étudiant est censé connaître. Nous n'essaierons donc pas ici de répondre à la question « Qu'est-ce que connaître un mot? » ni de comparer dans le détail les types de compétences mesurés par les différents tests.

Comme l'indique notre titre, nous tenterons de déterminer quels mots les étudiants doivent connaître à leur entrée à l'université en examinant tous les mots sur lesquels portent les questions de l'examen choisi. Ces mots seront considérés comme des échantillons d'un ensemble dont les contours seront précisés par leur regroupement en champs lexicaux et par une confrontation systématique avec des ouvrages de référence, vocabulaires fondamentaux, listes de fréquence, manuels d'apprentissage, etc.

Au printemps 1992, J. Bourguignon et A. Charbonneau de l'Université de Montréal ainsi que R.-L. Roy et F. Lépine de l'Université Laval ont terminé l'élaboration du « Test de français Laval-Montréal » (désormais: TFLM). Ce test à correction objective avait été créé dans le but de doter ces deux universités d'un instrument de mesure de la connaissance du français des étudiants universitaires.

Cet examen, qui a été soumis à différents tests de validité, existe en six versions dont les cinq premières ont été analysées pour les besoins de cet article[1].

Le TFLM se divise en cinq parties, qui portent sur l'orthographe d'usage, l'orthographe grammaticale, la morphologie, la syntaxe et le vocabulaire. Chaque version comprend 19 questions sur

[1] Il n'a pas été possible de consulter la sixième version pour cette première étude.

le vocabulaire dont neuf portent sur la connaissance du sens des mots. Pour cette première analyse, nous n'avons conservé que ces neuf questions pour chacun des cinq examens, ce qui nous donne un corpus de 45 mots.[2]

**Caractérisation du vocabulaire visé par le TFLM**

Nous nous sommes interrogées en premier lieu sur le degré de spécialisation du vocabulaire visé par le TFLM. Comme ce test est conçu pour les étudiants qui commencent le premier cycle universitaire, on peut faire l'hypothèse qu'il ne suppose pas connus les termes spécialisés des disciplines ou champs d'études qu'on enseigne à l'université, comme dendrite, terme d'anatomie désignant les ramifications des cellules nerveuses.

En revanche, puisqu'il s'agit d'évaluer des étudiants qui ont fait deux ans d'études générales post-secondaires, on peut supposer que pourra faire l'objet d'une question un mot ou un emploi qui, sans appartenir à une terminologie donnée, n'existe que dans la langue savante (ouvrages pédagogiques, etc.) et non dans la langue parlée ordinaire. Autrement dit, nous nous attendons à trouver dans le test des mots que le *Grand Robert* 1985 qualifie de « didactiques » ou de « scientifiques » sans mentionner particu-lièrement une science ou une technique particulière, comme densimétrie.

La première hypothèse est confirmée: un seul des mots faisant l'objet d'une question appartient à une terminologie spécialisée, celle de la philosophie. Mais la seconde hypothèse ne l'est pas: aucun des mots n'est qualifié de didactique ou de scientifique dans ce dictionnaire. Notons cependant que trois des mots du corpus y sont qualifiés de littéraires.

---

[2] Ce test étant toujours utilisé par les Universités Laval et de Montréal, il ne nous est pas possible de dévoiler les mots qui composent ce corpus; les personnes intéressées peuvent cependant communiquer avec nous.

On peut donc qualifier le lexique visé par le TFLM de général, par opposition à spécialisé, et de courant, par opposition à savant.

En deuxième lieu, nous nous sommes penchées sur la fréquence des mots du test. La fréquence des mots est considérée par les didacticiens comme un indice d'utilité très important. J. Picoche, dans son ouvrage récent *Didactique du vocabulaire français* (1993) rappelle que l'on couvre 90% des occurrences de n'importe quel texte avec moins de 1 000 mots et a retenu la fréquence comme critère d'inclusion des mots dans un dictionnaire d'apprentissage systématique qu'elle est en train de préparer.

Cependant, on en conviendra facilement, la fréquence n'est pas un critère absolu puisque ce qui fait le contenu sémantique d'un texte est apporté précisément par les mots liés aux thèmes qu'il aborde, et qui peuvent avoir une fréquence très basse. En outre, les mots les plus fréquents sont connus de tous les francophones et ne sont intéressants pédagogiquement que pour l'enseignement du français langue étrangère.

Pour qualifier de ce point de vue le vocabulaire visé par le TFLM, nous avons d'une part, vérifié si les mots sur lesquels on a interrogé les étudiants appartiennent au vocabulaire fondamental et, d'autre part, nous avons déterminé leur rang dans le *Dictionnaire des fréquences* de P. Imbs *et al.* (1971).

Les mots qui font l'objet de questions dans le TFLM n'appartiennent pas au vocabulaire le plus fréquent dans la langue parlée: ils ne figurent pas dans les listes du *Français fondamental* (s.d.) premier et second niveau, et seulement deux d'entre eux figurent dans le *Dictionnaire des mots essentiels* de G. Matoré (1963).

Ils n'appartiennent pas non plus au vocabulaire le plus fréquent et le plus largement répandu de l'expression et de la pensée scientifique, vocabulaire défini par le CRÉDIF en 1971 et appelé *Vocabulaire général d'orientation scientifique* ou VGOS. Seulement deux des mots du corpus figurent dans ce répertoire.

On ne peut cependant qualifier les mots du TFLM de rares puisqu'ils se situent en moyenne au 10 000$^{e}$ rang ( sur 70 000) du *Dictionnaire des fréquences* qui, rappelons-le, a été construit à partir

d'un corpus de textes écrits, surtout littéraires, des XIX^e^ et XX^e^ siècles. Aucun des mots du TFLM n'a un rang supérieur à 28 772 et trois seulement ont un rang supérieur à 12 000.

On peut donc conclure que les mots que le TFLM suppose connus sont de fréquence intermédiaire, voire relativement élevée dans la langue écrite.

En troisième lieu, nous nous sommes demandé si le vocabulaire qui faisait l'objet d'évaluation était facile ou difficile. Il ne viendrait à l'idée de personne d'interroger des étudiants de niveau universitaire à propos de mots comme table ou aimer, qui ne présentent aucune difficulté particulière. Mais ne les évaluer que sur les chausse-trappes du pluriel de dilettante ne serait guère accepté par les temps qui courent.

Il nous a paru intéressant de vérifier la difficulté des mots retenus par le test que nous avons choisi. Nous considérons que leur présence dans le *Nouveau dictionnaire des difficultés du français moderne* de J. Hanse et D. Blampain (1994) révèle un degré de difficulté supérieur. L'examen montre qu'un peu moins du quart des mots fait l'objet d'une remarque dans ce répertoire. Les difficultés qui y sont évoquées pour les mots du TFLM sont de nature variée: existence d'un homonyme, d'un paronyme, morphologie irrégulière, etc.

On peut donc conclure que le lexique évalué ne se limite pas aux mots faciles mais qu'il n'est pas d'une extrême difficulté.

Dans un quatrième temps, nous avons voulu déterminer la variété de langue à laquelle appartenaient les mots du TFLM. On peut considérer que la difficulté de certains mots est intrinsèque, tandis que d'autres peuvent poser des problèmes aux étudiants à cause de la variation régionale et sociale, d'une part, et de l'existence d'emprunts non acceptés par la norme, d'autre part.

Nous avons donc également vérifié la présence d'anglicismes, de régionalismes et de mots populaires et vulgaires dans le TFLM. Il ne comporte aucun mot de ce type et, en outre, il ne comporte aucun mot qui en serait synonyme. Tous les mots du test sont standards. Ils le sont non pas en vertu d'une norme qui proscrirait leurs concurrents,

mais parce que le lexique-cible de ce test échappe à ces types de variation, dont on sait qu'ils concernent surtout l'usage quotidien et oral de la langue. Signalons toutefois que des tests utilisés précédemment à l'Université Laval comprenaient des questions portant sur les « diverses incorrections du vocabulaire », parmi lesquelles figuraient des régionalismes et des anglicismes, ce qui indique que l'inclusion de mots de ce type peut, aux yeux des spécialistes, présenter une certain intérêt.

La cinquième étape de notre analyse concerne le contenu sémantique des mots du TFLM. À quels champs lexicaux les mots qui ont servi à évaluer la compétence lexicale des étudiants appartiennent-ils? Peut-on voir une certaine unité sémantique entre ces dizaines de mots apparemment disparates?

Lorsqu'on examine ce vocabulaire du point de vue des catégories générales de sens, on constate aisément qu'il s'agit pour l'essentiel d'un vocabulaire abstrait qui désigne des qualités, des relations, des opérations ou des procès qui ne sont pas de nature sensible. Ce caractère abstrait est confirmé par la morphologie: plus de la moitié des substantifs sont construits avec les suffixes -ion, -tion et -ation.

Lorsqu'on classe ces mots en fonction de leur champ lexical, on fait les constatations suivantes: presque tous les verbes désignent des opérations de la pensée ou du discours, l'action du logos; il en est de même de plus de la moitié des substantifs, la plupart des adjectifs peuvent qualifier des opérations ou des démarches intellectuelles ou les discours qui les expriment; d'autres adjectifs, très nombreux aussi, désignent des qualités générales; d'autres, plus rares, désignent des dispositions ou des qualités morales. On peut donc conclure en disant qu'il s'agit d'un vocabulaire abstrait, centré sur la vie intellectuelle.

Enfin, nous avons étudié les mots du TFLM du point de vue de leur origine. Les mots nécessaires à l'expression de la vie intellectuelle sont normalement aussi anciens dans une langue que cette vie intellectuelle elle-même. Mais l'histoire particulière du français et sa longue coexistence avec le latin expliquent que ce n'est pas le cas pour notre langue et donnent un caractère double à notre vocabulaire. Nous

ne rappellerons pas ici les raisons qui expliquent la distinction des historiens de la langue entre, d'une part, les mots populaires, ayant été soumis à l'évolution phonétique –et dans bien des cas ayant connu d'importants changements sémantiques– et formant ce qu'on appelle le fonds primitif latin de notre vocabulaire et, d'autre part, les mots empruntés au latin ou au grec, surtout à partir du moyen français, et n'ayant subi qu'une francisation superficielle.

Il est particulièrement remarquable qu'un dixième seulement des mots choisis par les auteurs du test que nous décrivons appartient au fonds primitif du lexique français. Outre ces mots, un est qualifié par les dictionnaires historiques de demi-savant et un est un emprunt à l'italien. Tous les autres mots sont savants. Ils sont en grande majorité d'origine latine, mais quelques-uns sont d'origine grecque.

**Comparaison du TFLM avec des tests français et belges**

Animés du même souci que les auteurs du TFLM, des universitaires d'autres pays francophones ont tenté de mesurer la compétence de jeunes qui se présentent au seuil de l'université et de déterminer si elle leur permet d'espérer le succès dans ces études.

À Liège, Ch. Delcourt et J. Delcourt-Angélique ont élaboré un test de français en 1979 dont 20 questions portaient sur le vocabulaire. Les mots choisis l'ont été parmi les 7 000 vocables les plus fréquents de la langue française. Les catégories sémantiques sont le concret, le psychologique, le didactique et l'abstrait.

À Louvain, J. Klein et J.-M. Pierret ont également élaboré un test de français qui comporte 25 questions sur le lexique. Les mots ont été sélectionnés en fonction de leur appartenance à la langue du discours didactique de portée générale et non scientifique. Ils ont puisé ces mots parmi trois catégories de vocables: des mots d'un niveau assez soutenu, des mots faisant partie d'un vocabulaire didactique que l'on pourrait qualifier d'opérateurs méthodologiques et des termes scientifiques plus ou moins vulgarisés. Selon eux, ces mots sont le reflet d'un niveau de compétence indispensable pour aborder un discours rationnel, argumenté, de type scientifique ou didactique du

genre de celui qui se rencontre dans les cours de la première année universitaire.

À Rennes, en 1982, Cl. Lebrun a préparé et administré à des classes de seconde un test de français de cinquante questions, dont la moitié portait sur le vocabulaire. Elle a choisi les mots devant faire l'objet d'évaluation en fonction des critères suivants: ils devaient faire partie du vocabulaire usuel nécessaire pour tirer profit d'un manuel d'histoire ou de géographie, ils devaient permettre de saisir le sens d'un texte littéraire simple et de s'exprimer à propos de ce texte et révéler la maîtrise de notions indispensables au travail intellectuel.

On le voit, les descriptions du lexique à tester sont fort semblables. Cette similitude s'explique par la commune nature des compétences nécessaires et par le sentiment partagé, nouveau et simultané qu'elles ne sont peut-être plus suffisamment répandues chez les étudiants qui devraient les posséder. Pour nous assurer qu'au-delà de ces descriptions générales, forcément approximatives, les auteurs des tests visaient *grosso modo* le même secteur du lexique, nous avons comparé systématiquement les mots retenus dans ces différents tests avec ceux du TFLM en nous plaçant successivement à tous les points de vue adoptés pour ceux-ci.

Le résultat de cette comparaison est très révélateur: partout, à Liège, à Louvain, à Rennes, le vocabulaire retenu présente les mêmes caractéristiques que celui du TFLM. La seule différence à noter est la plus grande proportion, dans le test de Louvain, de mots savants d'origine grecque. Cette convergence d'initiatives indépen-dantes les unes des autres nous paraît significative et est de nature à accroître la crédibilité de ces tests quant à la pertinence de la démarche entreprise par les auteurs et quant à la validité de leurs choix.

**Conclusion**

Nous conclurons cette étude –toute provisoire– par une rapide synthèse de nos résultats et par quelques réflexions inspirées surtout par la prise en considération de leurs conséquences pratiques.

En réponse à la question posée dans le titre de notre article, nous dirons qu'il est possible de circonscrire la compétence lexicale que les spécialistes estiment pouvoir exiger d'un étudiant qui commence des études universitaires.

La compétence évaluée par le TFLM et les tests auxquels nous l'avons comparé est la suivante: c'est la maîtrise d'un lexique général qui ne comprend pas de mots relevant de terminologies technique ou scientifique. Il n'est pas composé de mots propres à la langue savante, ce qui justifie qu'on le qualifie de courant. Il est composé de mots relativement ou moyennement fréquents dans la langue écrite, mais n'appartenant pas au vocabulaire de la langue parlée. Il inclut des mots considérés comme difficiles, mais la plus grande partie de ceux qui le composent ne présentent pas de difficulté particulière. C'est un lexique standard, qui n'est pas affecté par la variation sociale ni par la variation régionale et qui ne comprend pas de mots stylistiquement marqués. Il ne comprend pas non plus d'anglicismes. Il est essentiellement de nature abstraite et concerne surtout la vie intellectuelle. Historiquement, sa formation remonte à la relatinisation qui a marqué la fin du Moyen-Âge et la Renaissance: s'il n'est pas savant par son usage contemporain, ce lexique l'est par son origine.

L'ensemble ainsi circonscrit n'est pas disparate, mais constitue, au contraire, un tout homogène, un secteur bien limité du lexique français, aussi bien du point de vue synchronique que du point de vue historique.

Les caractéristiques que nous venons de mettre en évidence semblent indiquer que la vie courante ne fournit guère d'occasions d'employer le lexique qui semble, aux yeux des spécialistes de plusieurs pays, nécessaire à la réussite des études universitaires. Il est donc vain de penser que, mis à part quelques privilégiés particulièrement doués et motivés ou issus de milieux intellectuels, les étudiants peuvent l'acquérir par simple imprégnation. C'est donc à l'institution scolaire que revient la tâche d'assurer cette acquisition.

Or, au Québec, les documents officiels antérieurs à 1996 (par exemple, M.E.Q. 1981) ne prévoient pas d'activités centrées sur l'enseignement et l'apprentissage du vocabulaire –en particulier sur le

type de mots que nous avons décrit. Il nous semble que c'est là une lacune grave et qu'il faudrait sans tarder prendre des mesures pour corriger la situation.

Les mesures que nous préconisons sont d'autant plus nécessaires et urgentes que tout semble indiquer que le système scolaire actuel produit des cohortes d'étudiants dont l'ignorance en ce domaine est alarmante. C'est ainsi qu'un collègue belge, G. Legros, note qu'aux Facultés universitaires de Namur, « sur quelques 350 nouveaux étudiants, à peu près les trois quarts ignoraient la signification de virtuel, patent et corollaire et un peu plus de la moitié celle de empirique, hypertrophié, intrinsèque et éclectisme » (1995: 6). Selon une enquête faite à Bruxelles sur plus de 1 000 sujets, élèves des deux dernières années du secondaire et étudiants de la première année d'université, 34% de ces jeunes inversent les compléments dans substituer x à y (Dalcq, Van Raemdonck et Wilmet 1989: 10). Les constatations de Cl. Lebrun, à Rennes, vont dans le même sens: plus de la moitié des élèves de seconde ne connaissent pas le sens de éluder et près de 40% celui de réfuter (1992: 70).

De plus, plusieurs études ont démontré qu'il existe une corrélation très étroite entre la maîtrise de ce vocabulaire et la réussite scolaire (Lebrun 1992: 60-61, Dalcq, Van Raemdonck et Wilmet 1989: 10). Certes, de semblables corrélations existent entre les autres aspects de la compétence linguistique et la réussite scolaire, mais pour le lexique dont nous parlons, il est plus manifeste qu'ailleurs; plus que d'un simple lien statistique, il s'agit là d'une relation de causalité. En effet, les mots que l'étudiant est censé connaître, qu'il ne va plus de soi qu'il connaît et qu'effectivement il ne connaît pas toujours sont liés directement, organiquement, à la nature même de l'enseignement universitaire. On peut penser, à la rigueur, que la méconnaissance des règles d'accord du participe passé et de l'orthographe d'usage sont le plus souvent sans incidence directe sur la compréhension d'une démonstration scientifique, mais l'ignorance de mots comme tangible ou réfutation peut entraîner des contresens graves et compromettre l'acquisition d'un pan entier du savoir.

Ajoutons enfin qu'un enseignement systématique du vocabulaire général de l'abstraction compenserait en partie les effets de la non-connaissance du latin. Paradoxalement, cet enseignement a été supprimé au moment où il devenait le plus nécessaire, en même temps que l'on a supprimé celui du grec. Actuellement, les étudiants ne font ni grec ni latin et les mots abstraits, empruntés, pour la plupart, à ces langues, ne leur sont pas enseignés. Il y a quelques décennies, non seulement ils étaient couramment employés dans la classe de français, mais ils y faisaient en outre l'objet d'exercices répétés alors que, pourtant, la connaissance du latin et du grec, assurée par l'enseignement classique, aurait pu pallier en grande partie l'absence d'enseignement systématique des mots français savants ou de formation savante.

Cet enseignement devrait être systématique, car le contenu actuel de l'enseignement du français –notamment le choix des textes étudiés en classe– ne permet guère d'espérer qu'un nombre suffisant des mots dont nous parlons seront d'un emploi assez fréquent dans le cursus scolaire des élèves pour leur en permettre une acquisition spontanée et naturelle.

De même, la philosophie pédagogique de beaucoup de professeurs, orientée vers la sensibilisation de l'élève et valorisant la démarche qui amène le maître à se mettre au niveau de l'élève et non à prendre tous les moyens pour amener ce dernier au niveau visé n'est pas de nature à favoriser un accroissement du vocabulaire abstrait.

Bien entendu, cet enseignement devrait s'accompagner d'une évaluation de son efficacité. Il faudrait que l'on s'interroge à ce propos sur l'équivalence et sur la pertinence relative, en matière de vocabulaire, des tests dont nous venons de parler et de l'examen du M.E.Q., qui a remplacé au Québec les tests objectifs des universités. Nous pensons que cette équivalence n'existe pas et que la pertinence du test du M.E.Q. reste à démontrer pour la compétence lexicale décrite ici.

La nécessité d'un tel enseignement semble avoir été ressentie en Europe comme en fait foi la parution récente et presque simultanée de manuels consacrés à l'enseignement du vocabulaire, ou accordant

une place importante à cet enseignement. Il est remarquable que tous donnent une place centrale aux mots que nous avons cernés.

Au Canada, l'attention des lexicologues s'intéressant à la langue française s'est surtout concentrée sur quelques grands domaines: la terminologie, la description et la légitimation du fait régional, les études sur les emprunts, la didactique du vocabulaire aux très jeunes enfants ou aux non francophones et, plus récemment, les problèmes soulevés par la linguistique informatique. Ces choix semblent indiquer un intérêt louable pour des recherches liées au contexte social. Il nous semble que la transmission des outils lexicaux et conceptuels de la pensée abstraite à nos étudiants devrait retenir aussi leur attention.

Il faudrait, enfin et surtout, revaloriser l'abstrait, convaincre les « intervenants » du monde de l'éducation de la nécessité de l'abstraction dans la formation générale et persuader les décideurs de ne pas céder aux pressions démagogiques pour un enseignement encore plus concret. Beraud, Euzen-Dague et Rémi-Giraud (1988), auteurs du manuel d'enseignement du vocabulaire *Le Taste-mots dans les arbres*, notent que le vocabulaire qu'ils recensent est étranger à l'expérience vécue par leurs élèves, dont les préoccupations quotidiennes sont circonscrites et liées à des préoccupations essentiellement concrètes. C'est vrai, mais le rôle de l'école est précisément d'apporter à l'élève ce que son milieu immédiat ne peut pas toujours lui donner et de l'aider à dépasser par la pensée l'univers parfois étroit où le hasard l'a fait naître. Le monde se complexifie et le rôle de la science, domaine abstrait par excellence, s'accroît sans cesse; il ne faut pas que des déficiences linguistiques en écartent des jeunes qui pourraient y exceller avec une préparation adéquate.

**Bibliographie**

Beraud, A., M.-G. Euzen-Dague et S. Rémi-Giraud. 1988. *Le Taste-mots dans les arbres*. Lyon: CNDP.

CRÉDIF. 1971. *Vocabulaire général d'orientation scientifique*. Paris: Didier.

Dalcq, A.-E., D. Van Raemdonck et B. Wilmet. 1989. *Le français et les sciences*. Paris - Louvain-la-Neuve: Duculot.

France, Ministère de l'éducation - Direction de la coopération avec la Communauté et l'étranger. (s.d.). *Le Français fondamental*. Paris: Institut pédagogique national.

Hanse, J. et D. Blampain. 1994. *Nouveau dictionnaire des difficultés du français moderne*. 3e éd., Louvain-la-Neuve: DeBoeck-Duculot.

Imbs, P., R. Martin et R. Vienne (dir.). 1971. *Dictionnaire des fréquences: vocabulaire littéraire des XIXe et XXe siècles*. Paris: Didier.

Lebrun, Cl. 1992. « Pourquoi j'ai écrit *1000 mots pour réussir* ». *Enjeux*. 26: 56-82.

Legros, G. 1995. « Maîtriser le français écrit à l'université: un simple problème de langue? ». *Revue des sciences de l'éducation*. XXI, 1: 1-16.

Matoré, G. 1963. *Dictionnaire du vocabulaire essentiel*, Paris: Larousse

Ministère de l'Éducation. 1981. *Programme d'enseignement au secondaire. Français*. Québec: Éditeur officiel du.

Ministère de l'Éducation du Québec. 1992. *Guide d'évaluation d'un texte d'opinion*. Québec: Direction générale de l'évaluation et des ressources didactiques - Gouvernement du Québec.

Picoche, J. 1993. *Didactique du vocabulaire français*. Paris: Nathan.

Rey, Alain (dir.) . 1992. *Dictionnaire historique de la langue française*. Paris: Dictionnaires Le Robert.

Robert, Paul. 1985. *Le Grand Robert de la langue française. Dictionnaire alphabétique et analogique de la langue française*. 2e éd., Paris: Le Robert.

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 121-132.

# Lexicalisation disparate en milieu minoritaire

*Terry Nadasdi*
Université d'Alberta

## 1. *Introduction*

Dans cet article, je discute de l'emploi des pronoms objet direct en français ontarien (désormais, FO), en me concentrant notamment sur l'alternance entre les clitiques et les pronoms forts. Cette analyse tiendra compte d'un facteur linguistique, à savoir *la valeur évaluative du verbe*,[1] et d'un facteur social, la restriction dans l'emploi du français. L'hypothèse que je veux tester concerne l'emploi des clitiques parmi les locuteurs dont l'emploi du français est restreint (que j'appelle tout simplement les locuteurs restreints) et prévoit que ces locuteurs vont employer relativement moins de clitiques par rapport aux locuteurs non-restreints. Je m'intéresse également aux problèmes méthodologiques qui peuvent se manifester lorsqu'on examine la variation morphosyntaxique en milieu minoritaire. On verra par exemple qu'on peut parfois se heurter à des difficultés reliées à des associations lexicales qui n'existent que pour certains locuteurs et dont il faut absolument tenir compte afin de mesurer l'effet véritable de la restriction linguistique.

## 2. *Corpus*

Donnons d'abord quelques informations relatives au corpus que j'emploie pour mes analyses.[2] Les 120 locuteurs du corpus sont tous francophones et habitent quatre localités en Ontario, à savoir, Hawkesbury, Cornwall, North Bay et Pembroke. On peut dire que ce

---

[1]Pour une étude exhaustive de tous les facteurs qui conditionnent la variable, voir Nadasdi (1995a).

[2]Pour de plus amples renseignements sur ce corpus, voir Mougeon et Beniak (1991).

sont tous des francophones étant donné qu'ils ont au moins un parent francophone et qu'ils ont toujours fréquenté les écoles francophones. Toutefois, qu'il existe des différences importantes au sein de la communauté franco-ontarienne. Certains locuteurs emploient le français régulièrement dans la vie quotidienne, alors que d'autres ne l'emploient que très rarement. ce qui a des retombées importantes sur les structures qu'ils emploient et sur la fréquence d'emploi de ces structures. Donc un trait qui caractérise le français de beaucoup de Franco-Ontariens, c'est que leur français est "restreint" ou limité à certaines situations de communication. Afin de mesurer la restriction dans l'emploi du français, on a posé aux locuteurs du corpus plusieurs questions portant sur la fréquence de leur emploi du français et sur leur exposition à cette langue dans la communication de tous les jours. A partir des réponses à ces questions, les locuteurs ont été regroupés en trois catégories de restriction linguistique. Ces catégories sont données en (1).

(1) *Locuteurs à usage non restreint* (les LUNR)
*Locuteurs à usage moyennement restreint* (les LUMR)
*Locuteurs à usage fortement restreint* (les LUFR)

Il est à noter que même les locuteurs dont l'indice de restriction linguistique est très faible sont capables de s'exprimer en français, certes avec moins d'aisance que les locuteurs dont l'indice est élevé, mais avec assez de facilité pour soutenir une conversation en français. Ceci est sans doute largement imputable au fait que tous ces locuteurs ont bénéficié d'au moins neuf années de scolarisation dans une école de langue française. Les locuteurs à faible indice de fréquence d'emploi font donc montre d'une restriction dans l'emploi de la langue minoritaire qui est moins prononcée que celle d'autres locuteurs restreints examinés dans d'autres études sur les langues minoritaires (cf. Dorian, 1981, King, 1989).

### 3 *La variable*

La variable que j'examine ici est composée des *moyens pronominaux pour représenter un objet direct*. Cette variable se réalise sous la forme de deux variantes: a) un pronom faible qui se cliticise à la gauche du verbe et b) un pronom fort postposé au verbe. Des exemples de la variable sont présentés en (2), (3), et (4) où la variante clitique est donnée dans la série a).

(2) a. un areal ... on **l'**a pas mis sur la télévision. (Cor-05)
b. tous les bureaux on a tourné **ça**. (Pem-08)

(3) a. ce char ... il a été **le** chercher. (Nor-35)
b. un stéréo ... j'ai acheté **ça**. (Cor-11)

(4) a. voler de l'argent ... ils **l'**ont fait. (Nor-34)
b. parler anglais ... on est habitué à faire **ça**. (Pem-12)

Ces exemples révèlent que le système pronominal du FO s'écarte de celui de la langue standard en ce que la forme *ça* ne se limite pas à l'emploi démonstratif, mais fait aussi partie des pronoms objet direct, tout comme c'est le cas en français québécois. L'idée que *ça* fait partie des pronoms objet trouve de l'appui dans Thibault (1983: 59), où il est proposé qu'en français montréalais ce pronom, tel qu'il est employé dans les exemples de (2) à (4), "est un simple représentant qui n'a pas de valeur démonstrative" et qu'il fait donc partie du paradigme des pronoms anaphores, ce qui permet de le considérer comme une variante du clitique objet direct (désormais, CLOD).

Un autre trait du FO qu'on ne retrouve pas en français standard est la possibilité d'avoir un pronom fort objet direct autre que *ça* qui représente un référent animé.[3] Des exemples de cet emploi se trouvent

[3] Nous verrons ci-dessous que l'emploi d'un pronom fort à référent humain se limite à un sous-groupe de locuteurs du FO, à savoir, ceux dont l'emploi du français est très réduit.

dans les phrases de (5) à (8).

(5) Chips il a payé **eux-autres**. (Nor-26)

(6) la pollution commence à détruire **nous-autres**. (Cor-11)

(7) Pis y'ont fait' une échange de personnages, tout' les pilotes qu'ont été perdus comme dans le "Bermuda Triangle", pis y'ont retourné **eux-autres** pis y'ont pris une vingtaine d'autres personnes avec eux pis... (Pem-24)

(8) Quand j' leur parle là, i'comprend **moi** en français. (Pem-16)

Dans ces exemples, les pronoms forts ne représentent pas un emploi déictique et ne portent pas d'accent d'insistance. D'après Thibault (1983: 60), cet emploi des pronoms forts ne se trouve pas en français montréalais où l'emploi d'un pronom fort [+animé] se limite à des constructions dans lesquelles cet élément connaît un emploi déictique. Il s'agit donc d'un trait qui distingue le système pronominal de ces deux variétés de français canadien.

4. *Résultats généraux sur les morphèmes liés*

Considérons d'abord pourquoi l'étude de l'alternance entre les clitiques et les pronoms forts dans un milieu minoritaire pourrait nous intéresser. Une première motivation serait les hypothèses de Andersen (1982) relatives aux morphèmes liés. D'après lui, les locuteurs restreints devraient faire montre d'une préférence marquée pour les structures où l'on emploie un morphème libre là où il existe un choix entre un morphème libre et un morphème lié. Si l'on accepte que les clitiques sont des affixes verbaux,[4] on a encore d'autres raisons de croire qu'on trouvera moins de clitiques parmi les locuteurs restreints,

[4]Cf. Roberge (1990), Auger (1994), Nadasdi (1995a).

car des études précédentes ont démontré que les affixes verbaux sont moins fréquents parmi les locuteurs restreints. Par exemple, dans une étude portant sur l'emploi des marques d'accord de la 3p. du pluriel des verbes irréguliers en FO, Mougeon et Beniak (1991) notent que les locuteurs restreints emploient ces affixes moins souvent que les locuteurs non restreints. Leurs résultats relativement au nivellement du marque de pluriel sont présentés en (9):

(9) locuteurs non-restreints: 19/1213 (2%)
locuteurs restreints: 203/1074 (19%)

Des résultats similaires sont présentés dans Dorian (1978), qui démontre que les semi-locuteurs du gaélique écossais, une langue en voie d'extinction, tendent à simplifier des structures morphologiques complexes.

Étant donnés ces résultats, on devrait s'attendre à ce que ces affixes verbaux qu'on appelle clitiques objet direct soient aussi des pierres d'achoppement pour les locuteurs restreints. En fait, d'autres études portant précisément sur les clitiques vont dans le sens d'une réduction dans l'emploi des clitiques parmi les locuteurs restreints. Par exemple, Nadasdi (1995a) examine l'alternance entre le clitique locatif *y* et la forme *là*, et conclut que la restriction linguistique a un effet très fort sur l'emploi de cette variable. Ces résultats sont donnés dans le Tableau 1:

**Tableau 1:** *Emploi du clitique locatif selon la restriction linguistique*

| **Restriction linguistique** | **Poids relatif** | **%** | **Occurrences** |
|---|---|---|---|
| LUNR | .791 | 69 | 87/127 |
| LUMR | .486 | 30 | 39/129 |
| LUFR | .188 | 7 | 8/111 |

Comme on peut le voir dans ce tableau, l'emploi du clitique locatif est

fortement conditionné par la restriction dans l'emploi du français, les locuteurs restreints n'employant les clitiques que dans 7% des cas, alors que les non restreints les emploient à raison de 69%.

Des études qui ont examiné les clitiques en espagnol ont trouvé des résultats qui vont dans le même sens. Par exemple, Silva-Corvalán (1994) note que les locuteurs dont l'emploi de l'espagnol est relativement restreint ont plus tendance à éviter les clitiques. Donc, pour ce qui est de l'alternance entre les clitiques objet direct et la forme *ça*, on devrait s'attendre à ce que les LUFR emploient relativement plus de *ça*.

Passons donc aux résultats relatifs à l'effet la restriction linguistique sur les clitiques objet directs. Ces résultats sont présentés dans le Tableau 2.

**Tableau 2:** *Effet de la restriction linguistique sur l'emploi des clitiques objet direct*

| **Restriction linguistique** | **Poids relatif** | **%** | **Occurrences** |
|---|---|---|---|
| LUNR | .586 | 74 | 762/1028 |
| LUMR/LUFR[5] | .456 | 76 | 1531/2007 |

La première chose qu'on note dans le Tableau ici, c'est que la restriction linguistique ne semble avoir qu'un effet très modeste sur la variable. En effet, si l'on examine le Tableau 2, on voit qu'en termes de pourcentages, les LUFR emploient plus de clitiques que les locuteurs non restreints! Ce résultat est tout à fait le contraire de ce que prévoit mon hypothèse de départ et ne s'accorde absolument pas avec les résultats issus d'autres études sur l'emploi des morphèmes liés dans les langues minoritaires.

Je crois qu'on peut proposer deux explications pour les

[5]La distinction entre les LUMR et les LUFR n'est pas significative pour cette variable.

résultats du Tableau 2: a) la tendance des locuteurs restreints à employer moins de clitiques existe mais elle est manifestement très faible; b) la restriction linguistique est en interaction avec d'autres facteurs et cela a pour effet de dissimuler partiellement son effet réel (cf. Poplack, sous presse). Afin de tester cette deuxième possibilité, je voudrais considérer l'effet d'un des groupes de facteurs linguistiques qui conditionnent la variable, à savoir *la valeur évaluative du verbe*.

5.1. *Les verbes à valeur évaluative*

Le groupe de facteurs *valeur évaluative du verbe* comprend deux facteurs: a) les verbes évaluatifs b) les verbes non évaluatifs. Dans cette première catégorie, on retrouve les verbes comme *détester*, *adorer*, et notamment *aimer*. Il s'agit de verbes dont le sujet porte un jugement de valeur sur l'objet direct. Dans la deuxième catégorie se trouvent les autres verbes, c'est-à-dire ceux dont le sujet ne porte aucun jugement sur l'objet verbal. Des exemples de la variable avec ce groupe de facteurs sont donnés en (10) et (11):

**verbe évaluatif**

(10) a. disons qu'i' visit'rait cet hôtel-là... pis il **la** trouve belle hein... (Cor-23)
b. j'aime bien **ça** comme ... des aventures. (Haw-12)

**verbe non évaluatif**

(11) a. La famille du Québec ... je l'ai pas vu. (Nor-18)
b. Les Tannants ... je regarde **ça**. (Cor-20)

L'hypothèse que je voulais tester avec ce groupe de facteurs est que lorsqu'un locuteur emploie un verbe évaluatif, il aura tendance à se servir d'un pronom fort et postposé, c'est-à-dire séparé du verbe, de sorte que cet élément soit mis en relief, non pas par un accent d'insistance, car les objets directs portant un accent d'insistance ont été écartés de mes analyses, mais par le simple fait de ne pas être intégré à la morphologie verbale, contrairement à la variante clitique (cf.

Thibault, 1983). Donc, on s'attend à ce que la présence d'un verbe évaluatif soit défavorable à l'emploi des clitiques.

Examinons donc le Tableau 3 où l'on trouve les résultats relatifs à ce groupe de facteurs:

**Tableau 3:** *Effet des verbes évaluatifs sur l'emploi des clitiques objet direct*

| Type de verbe | Poids relatif | % | Occurrences |
|---|---|---|---|
| non évaluatif | .604 | 85 | 2207/2603 |
| évaluatif | .073 | 20 | 86/432 |

Tout comme le note Thibault (1983) dans son étude des pronoms objet en français montréalais, les verbes évaluatifs défavorisent nettement l'emploi des CLOD en FO. Et, comme je l'ai déjà suggéré, cela peut s'attribuer au fait que l'emploi d'un pronom fort permet de mettre en relief l'objet direct de cette classe de verbes par le simple fait d'être séparé du verbe. L'emploi très élevé de *ça* avec les verbes évaluatifs, qui est dans la grande majorité des cas le verbe *aimer*, suggère qu'une suite comme *j'aime ça* est en quelque sorte lexicalisée pour former une expression figée. Tout au moins on peut dire qu'il existe une association lexicale très forte entre la variante *ça* et *aimer*. Mais est-ce que cette affinité est la même pour tous les locuteurs du corpus? Normalement, il faut donner une réponse positive à cette question étant donné que l'effet d'un groupe de facteurs linguistique est censé être plus ou moins identique pour tous les membres de la communauté. Toutefois, cela n'est pas forcément le cas pour les langues parlées en milieu minoritaire (cf. Mougeon et Nadasdi, 1996).

### 5.2. *Verbes évaluatifs et restriction linguistique*

On vient de voir qu'il existe une association lexicale très forte entre la variante postposée *ça* et les verbes comme *aimer*. Toutefois, il se peut que parmi les locuteurs dont la restriction dans l'emploi du

français est assez élevée, l'affinité entre ces formes soit moins prononcée et, donc, que même si en général les pronoms forts sont employés plus souvent par ces derniers locuteurs, cela ne soit pas le cas après les verbes évaluatifs. Cela me semble tout à fait plausible si d'une part l'association lexicalisation résulte d'un emploi fréquent de deux formes et que d'autre part les LUFR emploient moins souvent le français. J'ai donc décidé d'examiner l'effet de la valeur évaluative du verbe séparément pour chaque catégorie de restriction. Le Tableau 4 présente les résultats d'une comparaison de l'emploi de *ça* avec les verbes évaluatifs selon les trois niveaux de restriction linguistique:

**Tableau 4:** *Influence combinée des groupes de facteurs "verbe évaluatif" et "restriction linguistique" sur l'emploi de ça objet direct*

| Restriction linguistique | Occurrences | % |
|---|---|---|
| LUNR | 184/206 | 89 |
| LUMR | 114/146 | 78 |
| LUFR | 48/80 | 60 |

Comme on peut le constater, l'emploi de *ça* avec les verbes évaluatifs est nettement plus basse parmi les locuteurs dont l'emploi du français est réduit.[6] Le taux moyen des pronoms forts est de 85% avec les verbes évaluatifs, mais cela descend à 60% pour les locuteurs dont l'emploi du français est fortement restreint. Il est donc évident que l'association lexicale de *ça* avec cette classe de verbes est moins forte chez les LUFR que chez les LUNR. On se trouve donc en face de deux tendances opposées: a) la tendance des locuteurs restreints à employer moins de clitiques en général et b) la tendance de ces mêmes locuteurs à employer relativement plus de clitiques avec les verbes évaluatifs.

[6]Notons que tout comme les locuteurs de Thibault, les LUNR emploient 89% de *ça* avec les verbes évaluatifs.

C'est parce que ces tendances vont à l'encontre l'une de l'autre que l'effet de la restriction linguistique sur l'emploi des clitiques objet direct ne se manifeste que de façon relativement modeste. Il s'ensuit que si l'on fait abstraction des occurrences de la variable qui contiennent un verbe évaluatif, l'effet de la restriction linguistique ou bien celui de la localité devrait être plus net. Cela est confirmé dans le Tableau 5 où je donne les poids relatifs pour la restriction linguistique en excluant les occurrences où l'on emploie un verbe évaluatif.

**Tableau 5:** *Effet de la restriction linguistique sur l'emploi des clitiques objet direct en faisant abstraction des verbes évaluatifs*

| **Restriction linguistique** | **Poids relatif** | **%** | **Occurrences** |
|---|---|---|---|
| LUNR | .657 | 90 | 740/822 |
| LUMR+LUFR[7] | .426 | 83 | 1467/1781 |

Il est vrai que les différences entre les résultats de ce tableau et ceux du Tableau 2 ne sont pas énormes, mais l'effet de la restriction linguistique est nettement plus fort et ce résultat est bien plus conforme à nos attentes que ce qu'on a noté dans notre première analyse.

6. *Conclusion*

Pour conclure, le problème qu'on a ici, c'est qu'il y a de l'interaction entre deux groupes de facteurs: *la valeur évaluative du verbe* et *la restriction dans l'emploi du français*. Comme le montrent Mougeon et Nadasdi (1996), les interactions de ce type sont assez communes dans les communautés linguistiques minoritaires et elles peuvent se manifester de plusieurs manières. Cela se manifeste lorsqu'il

[7]Signalons que l'exclusion des verbes évaluatifs ne nous permet pas de différencier les LUMR des LUFR car l'écart entre ces deux groupes n'est toujours pas significatif (LUMR = .431; LUFR = .420).

y a une variante qui n'est employée que par un sous-groupe de locuteurs; cela se réalise aussi lorsque l'effet relatif d'un groupe de facteurs linguistique est différent pour différents sous-groupes de locuteurs, ce qui est le cas pour la variable que je viens de présenter. On a vu, par exemple que l'effet de la valeur évaluative du verbe n'est pas le même pour les trois groupes de locuteurs qu'on peut distinguer selon leur emploi du français. Un troisième type d'interaction qu'on peut trouver dans les communautés minoritaires se manifeste lorsque deux groupes de facteurs sociaux sont en interaction. Je discute d'un cas de ce type dans Nadasdi (1995b) où je démontre que la classe socio-économique n'est pertinente que pour les locuteurs non restreints dans l'alternance entre les sujets redoublés et non redoublés. L'existence des interactions de ce type fait en sorte que toute étude qui cherche à mesurer l'effet véritable de la restriction linguistique doit en tenir compte.

**Bibliographie**

Andersen, Roger. 1982. "Determining the Linguistic Attributes of Language Attrition". *The Loss of Language Skills*, R. Lambert & B. Freed (dirs.). Rowley, Mass.: Newbury House. 83-118.

Auger, Julie. 1994. *Pronominal Clitics in Québec Colloquial French: A Morphological Analysis*. Thèse de doctorat, University of Pennsylvania. (Distribuée par IRCS, Rapport 94-29)

Dorian, Nancy. 1978. "The Fate of Morphological Complexity in Language Death: Evidence from East Sutherland Gaelic". *Language* 54. 590-609.

Dorian, Nancy. 1981. *Language Death: The Life Cycle of a Scottish Gaelic Dialect*. Philadelphia: University of Pennsylvania Press.

King, Ruth. 1989. "Le français terre-neuvien: aperçu général". *Le français canadien parlé hors Québec: aperçu sociolinguistique*, R. Mougeon & É. Beniak (dirs.). Québec: Presses de l'Université Laval. 227-244.

Mougeon, Raymond & Édouard Beniak. 1991. *The Linguistic Consequences of Language Contact and Restriction: the Case*

*of French in Ontario, Canada*. Oxford: Oxford University Press.

Mougeon, Raymond & Terry Nadasdi. 1996. "Sociolinguistic Discontinuities in Minority Language Communities". Ms., York University.

Nadasdi, Terry. 1995a. *Variation morphosyntaxique et langue minoritaire: le cas du français ontarien*. Thèse de doctorat, University of Toronto.

Nadasdi, Terry. 1995b. "Subject NP doubling, matching and minority French". *Language Variation and Change* 7: 1. 1-14.

Poplack, Shana. sous presse. "The Sociolinguistic Dynamics of Apparent Convergence". *Towards a Social Science of Language. Papers in Honor of William Labov. Volume 2: Social Interaction and Discourse Structures*, G. Guy, Gregory R. C. Feagin, D. Schiffrin, & J. Baugh. (dirs.). Amsterdam: Benjamins.

Roberge, Yves. 1990. *The Syntactic Recoverability of Null Arguments*. Montréal: McGill-Queen's University Press.

Silva-Corvalán, Carmen. 1994. *Language Contact and Change: Spanish in Los Angeles*. Oxford: Oxford University Press.

Thibault, Pierrette. 1983. *Équivalence et grammaticalisation*. Thèse de doctorat, Université de Montréal.

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 133-145.

# Deux champs sémantiques chez les anglophones et les francophones de Montréal[1]

*Gillian Sankoff*
University of Pennsylania

## 1. Introduction

Ceux qui étudient le bilinguisme ont souvent fait la remarque que les locuteurs d'une langue seconde, même ceux qui sont très compétents, possèdent des répertoires moins différenciés que les répertoires qui leur sont disponibles dans leurs langues maternelles. Il y a cependant peu d'évidence à l'appui d'une telle assertion. Cette situation n'est pas surprenante quand on considère l'ampleur de la recherche nécessaire pour la vérifier. Dans le domaine de la variation phonologique, par exemple, il faudrait disposer d'une étude de la gamme de variation stylistique chez les locuteurs de langue maternelle, pour ensuite mener une étude parallèle auprès des locuteurs de langue seconde, préférablement chez ceux qui auraient comme cible le modèle de cette même communauté[2].

La recherche que j'ai entreprise avec Pierrette Thibault en 1993 a pour principal objectif d'établir dans quelle mesure les anglophones de Montréal ont réussi à s'intégrer au sein de la communauté linguistique francophone de cette ville. Pour y arriver, nous comparons la production linguistique des anglophones à celle des locuteurs natifs, en utilisant des données tirées des corpus Sankoff-Cedergren 1971 (Sankoff et al. 1976) et Montréal 1984 (Thibault et Vincent 1990) dont nous disposons.

Le présent article vise à comparer l'usage d'un échantillon d'anglophones montréalais avec celui des francophones montréalais en ce qui concerne le lexique dans deux champs sémantiques: les verbes qui se rapportent au domaine de la "résidence", et les noms référant au "travail". Il s'agit de deux domaines qui font partie du vocabulaire de base, et non pas de termes spécialisés ou recherchés. Même des locuteurs dont le niveau de compétence en langue seconde est faible trouveront des façons de parler d'où ils habitent et du

[1] Je voudrais remercier tout particulièrement ma collaboratrice Pierrette Thibault, qui m'a aidée en discutant plusieurs aspects de l'analyse et dans la préparation du texte pour publication.

[2] Archambault et al. (1996) constitue un pas dans cette direction. Même si la variation stylistique comme telle n'y est pas étudiée, le répertoire phonétique des locuteurs de L2 est comparé à celui des locuteurs de L1 en ce qui concerne plusieurs caractéristiques du français québécois.

BEST COPY AVAILABLE

travail qu'ils font. Cependant, si l'on se fie à la littérature sur le bilinguisme, on s'attendra à ce que les locuteurs de langue seconde aient moins de mots disponibles et qu'ils les utilisent avec moins de subtilité sémantique et stylistique que les locuteurs natifs.

C'est en effet ce que tendent à montrer deux études entre autres. D'abord, dans une expérience en psycholinguistique, Lambert et al. (1959) évalue le degré de bilinguisme à partir du nombre de synonymes facilement trouvés par un locuteur. Pour exprimer l'état d'avoir trop bu, les locuteurs natifs trouvent rapidement non seulement le mot livresque, *ivre*, mais aussi des synonymes moins formels comme *saoûl*, *paqueté*, etc. De plus, notre propre étude des marqueurs discursifs utilisés par les anglophones montréalais en français (Sankoff et al. 1997) tend aussi à montrer une plus grande disponibilité de termes chez les locuteurs les plus compétents.

Les données présentées ici sont tirées des interviews sociolinguistiques de 29 locuteurs, enregistrées à Montréal entre 1993 et 1995. Les locuteurs natifs sont ceux du corpus Sankoff-Cedergren de 1971. Tous les locuteurs de langue seconde sont nés entre 1960 et 1975: au moment de l'enregistrement, ils avaient entre 18 et 33 ans. Bref, il s'agit de la génération d'anglophones qui a eu la possibilité de profiter des classes d' "immersion" en français au sein des écoles publiques anglophones. Bien sûr, nous avons tâché de représenter non seulement le groupe de jeunes adultes "post-immersion", mais aussi ceux qui ont eu moins ou davantage de contacts avec le francais, et à l'école et dans leurs milieux sociaux, durant leur enfance et leur adolescence. Parmi les 29, 27 se considèrent comme anglophones; les deux autres à qui on a attribué les pseudonymes "Vincent" et "Jocelyne", se considèrent comme "bilingues"[3]. Nous classons ces deux locuteurs dans une catégorie séparée, pour mieux cerner ce qui les distingue des autres.

Les modèles d'usage de mots pour tous ces locuteurs dans les deux champs sémantiques en question seront comparés à ceux

[3] Jocelyne a grandi dans une famille où les deux parents étaient Québécois francophones, mais se déclare plus à l'aise en anglais maintenant, au point où elle essaie de ne pas révéler son nom de famille (un nom typiquement québécois) à son travail parce qu'elle pense que les gens vont dire "Oh la pauvre fille, elle a perdu son français". Vincent, par contre, est fils d'une famille immigrante anglophone mais il a été adopté à l'âge de 2 ans par une famille francophone montréalaise. Il a pourtant gardé des liens avec sa famille d'origine.

1985

des locuteurs natifs analysés par D. Sankoff, Thibault et Bérubé (1978) [dorénavant désignés par STB].

## 2. Méthodologie

La méthodologie élaborée par STB a été suivie de près. Dans les deux cas, la recherche prend appui sur la production linguistique spontanée qui caractérise les interviews sociolinguistiques. Comme STB, j'ai repéré chaque occurrence de tous les verbes employés dans les deux champs sémantiques retenus, tout en prenant soin d'exclure les cas où l'interviewé aurait pu être influencé par le choix de mot de l'intervieweur. Ainsi des exemples comme (1) dans des énoncés contigus ont été exclus, alors qu'ont été conservés les cas où le locuteur choisit un mot sans subir l'influence de l'intervieweur, ou lorsqu'il emploie un mot différent de celui-ci comme dans (2).

(1) Int: Est-ce qu'ils habitent aussi dans le coin?
Sandra: Oui bien mes parents habitent aussi sur l'avenue du Parc.
(2) Int: Elle réside plus ici, c'est ça?
Karl: Non non, elle vit à Dorval.

Dans l'analyse des traits sémantiques, j'ai également repris les catégories de STB, tel qu'expliqué dans la section qui suit.

## 3. Verbes de 'résidence'

### 3.1 Définition du champ sémantique

Selon le modèle proposé dans STB, les champs sémantiques se définissent par la co-présence d'un certain nombre de traits qu'ils peuvent ou non avoir en commun. En général, ces traits tendent à se chevaucher partiellement.

En ce qui concerne les verbes de 'résidence', nous remarquons que plusieurs parmi ceux qui véhiculent cette notion sont aussi employés dans des sens moins spécifiques. Par exemple, le verbe *vivre* peut s'employer pour parler tout simplement de l'expérience de vivre, sans impliquer la notion de résidence. Selon le schéma de STB, *vivre* véhicule alors uniquement le trait d'existence, symbolisé par la lettre *e*. *Vivre* est d'ailleurs le seul verbe utilisé dans ce sens, comme dans l'exemple tiré de mon échantillon.

Sens A, *e* [+existence]: *vivre*
(3) J'ai appris comment est-ce-qu'on vit dans cette société bilingue. (Paul M.: 220[4])

[4] Nous avons attribué un prénom pseudonyme à chacun de nos locuteurs. Le chiffre qui suit le nom réfère à la ligne de la transcription.

Parmi les autres verbes qui signalent la notion de résidence se trouve *rester*. Tout comme *vivre*, *rester* peut aussi s'employer dans des contextes moins spécifiques. Il évoque alors une notion de 'manque de changement', symbolisée par *m*. Encore une fois, *rester* semble être le seul verbe à véhiculer ce sens.

Sens B. *m* [-changement]: *rester*

(4) Il en reste pas beaucoup -- la plupart sont morts. (Ted N:1171)

Un troisième cas combine le trait *m* avec un nouveau trait, celui de la continuité (*c*). Dans ce cas, il y a un autre verbe (*demeurer*) qui peut être utilisé, au moins chez les locuteurs natifs. Puisque nous n'en avons trouvé aucun cas chez les anglophones, nous incluons un exemple tiré du corpus Sankoff-Cedergren. Il faut souligner que même chez les locuteurs natifs, l'usage de *rester* dans ce sens est presque catégorique, et *demeurer* figure comme une option tout à fait marginale.

Sens C. *m* ∩ *c* [-changement, +continuité]: *rester ~ (demeurer)*

(5) Je restais après l'école pendant une heure ou une heure et demie. (Kathy K.:542)

(6) Je pense que la signification demeure toujours pas mal la même chose. (1971/25:548)

La notion de 'résidence' combine les trois traits déjà discutés, et elle peut s'exprimer par quatre verbes différentes: *vivre*, *rester*, *demeurer* et *habiter*[5] L'identification des verbes employés dans ce sens est facilitée par le fait qu'il sont toujours accompagnés d'un locatif, tel que nous pouvons l'observer dans les exemples (7) à (10).

Sens D. *m* ∩ *c* ∩ *e* [-changement,+continuité,+existence]: *rester ~ demeurer ~habiter ~ vivre*

(7) Je choisirais plutôt de vivre à Montréal que de vivre à Regina ou Edmonton. (Kathy K.)

(8) Ma mère, elle demeure là mais mon père il a récemment- il a eu un emploi à Manitoba- au Manitoba. (Sandra N:348)

(9) Dans notre famille mon frère il habite à Toronto. (Larry G:339)

(10) Pour même pas un an on a resté à St-Léonard, mais après ça on a toujours resté à Rosemont. (Joan O., 422-3)

---

[5] Ces quatre mêmes verbes s'emploient aussi pour exprimer un sens encore plus précis, celui de la cohabitation, mais nous ne nous attarderons pas à l'explorer dans le présent article, vu le nombre plutôt restreint d'occurrences dans nos données.

### 3.2 L'usage des anglophones et des francophones

Le Tableau 1 expose l'usage que font les Franco-montréalais des quatre verbes qui peuvent s'employer dans ce domaine. Il est évident qu'il y a une préférence marquée pour le verbe *rester*, qui est choisi dans presque les deux tiers des cas.

| Verbe | Nombre de cas | Pourcentage |
|---|---|---|
| *rester* | 537 | 64% |
| *demeurer* | 167 | 20% |
| *vivre* | 85 | 10% |
| *habiter* | 47 | 6% |
| Total | 836 | 100% |

Tableau 1. Le sens 'avoir sa résidence' ($m \cap c \cap e$) véhiculé par quatres verbes différents chez 60 locuteurs du corpus de Montréal 1971 (STB, p.29)

Avant de passer à l'usage que font les anglophones de ces verbes de résidence, il importe de mettre en parallèle les termes utilisés en français et en anglais pour véhiculer des sens connexes. Pour le sens A ci-dessus, la signification 'existence' ne permet qu'un seul verbe dans chaque langue: *vivre* en français et *live* en anglais. Pour les sens B et C, le français du Québec se sert presqu'uniquement du verbe *rester*. Ici, l'on dirait normalement *stay* en anglais, le verbe *remain* étant un choix plutôt limité à la langue écrite. Le Tableau 2 précise ces options dans les deux langues.

| Mot angl. | Sens | Mot fr. québ. | Autres mots possibles |
|---|---|---|---|
| ***live*** | A "expérience de vivre" | *vivre* | |
| ***live*** | D "avoir sa résidence" | ***rester*** | *vivre, habiter, demeurer* |
| *stay* | B,C "manque de changement; continuité" | ***rester*** | |

Tableau 2. Relations mots/sens en anglais et en français.

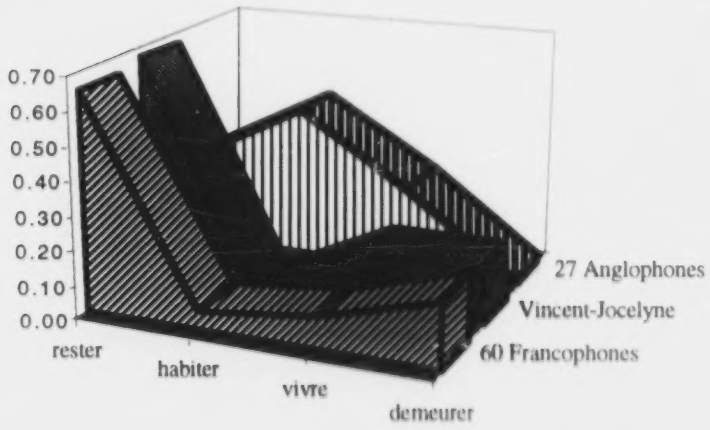
En ce qui concerne les sens D, la préférence des franco-montréalais pour *rester* (64%) est signalée par la place que ce verbe occupe dans le tableau. Chez les anglophones c'est *live*, et non *stay* qui est généralement employé. Chacune des deux langues distribue

deux verbes à travers ces traits, mais c'est uniquement pour le sens D qu'il y a manque de parallélisme.

Il y a donc lieu de se demander quel sera la préférence des anglophones pour véhiculer ce sens lorsqu'il parlent français? S'ils reproduisent le modèle de la communauté majoritaire francophone qui les entoure, ils utiliseront de préférence *rester*. S'ils se basent sur leurs connaissances de l'anglais, ils seront susceptibles de préférer l'usage de *vivre*, calqué sur l'anglais *live*. Il est évident que nous employons le terme "calque" ici non pas dans un sens absolu, puisque le verbe *vivre* est tout à fait légitime en français dans ce sens, et l'on n'y parlerait pas de calque dans l'usage d'un francophone. Chez les anglophones, il s'agirait plutôt d'une sur-utilisation par rapport au modèle des francophones, sur-utilisation qui résulterait de l'association établie par les bilingues entre *live* dans leur langue maternelle et *vivre* dans leur langue seconde. Si notre prédiction s'avère correcte, ce que Weinreich (1953:7) a appelé une "interlingual identification" entraînera un taux élevé de *vivre* au sens D.

Les résultats sont présentés dans le Graphique 1. Nos attentes n'y sont que partiellement confirmées. Il est vrai que les anglophones sur-utilisent *vivre* par rapport aux locuteurs natifs:

Graphique 1. Proportions relatives d'usage des verbes de "résidence" pour trois groupes de locuteurs



tandis que ces derniers affichent un taux de 10% d'usage de *vivre*, les anglophones l'utilisent plus de deux fois plus souvent, affichant un taux de 25%. Cependant, ce n'est pas leur choix préféré. En effet, ils utilisent à 46% le verbe *habiter*, verbe que les locuteurs natifs n'emploient que 6% du temps pour véhiculer le sens de résidence. Il s'agit du verbe qui, selon plusieurs ressortissants des cours de français dans les écoles anglophones, leur a été enseigné comme verbe approprié dans ce contexte.

Le Graphique 1 nous montre aussi les résultats pour les deux locuteurs bilingues, Vincent et Jocelyne. Leur usage de vivre (17%) est légèrement plus élevé que celui des locuteurs natifs; ils s'alignent toutefois sur ces derniers dans leur choix prépondérant de *rester* (64% pour les 60 Québécois francophones; 67% pour Vincent et Jocelyne), et dans leur utilisation de *demeurer* à 17% (20% pour les locuteurs natifs). Vincent et Jocelyne évitent complètement l'emploi d'*habiter*. Il est clair que du point de vue de la sémantique de ce champ, ces deux locuteurs font partie de la communauté francophone québécoise.

## 4. Noms référant au travail

### 4.1 Définition du champ sémantique

Dans le cas du champ de la résidence, nous avons étudié la distribution de plusieurs verbes. Dans le domaine du travail, il s'agit plutôt d'un groupe de substantifs. Le sens le plus large réfère à l'acte de travailler, et dans ce cas il est typiquement un nom de masse, s'employant souvent dans une construction partitive. Tandis que les anglophones n'utilisent que *travail* pour véhiculer ce sens, les locuteurs natifs emploient aussi les mots *ouvrage* et *job* (ce dernier ne figurant pas dans notre corpus, l'exemple (13) a été inventé).

A. *t* [+travail]: *travail ~ (ouvrage, job)*

(11) Le cours de statistiques ça prend beaucoup de travail.
(Glenda L: 318-9)

(12) Il y en a qui disent que l'ouvrage ne fait pas mourir.
(1971/94:281)

(13) Pelleter mon entrée de garage? C'est de la job!

Le deuxième sens ajoute au premier la notion d'un travail rémunéré, symbolisée par le trait *r*. Les noms *travail*, *emploi*, et *job* sont employés par les anglophones, les locuteurs natifs utilisant aussi le mot *ouvrage*.

B. *t* ∩ *r* [+travail, +rémunération]: *travail ~ emploi ~ job ~ (ouvrage)*

(14) J'avais un travail puis j'ai été slacké de ça juste avant le- le- le Noël. (Mike J:251)

(15) La compagnie a fermé. Alors il est à la recherche d'emploi mais ça va pas bien. (Kathy K:621)

(16) Après ça tu- ils trouvent toi une job dans la ma- pour l'été. (Ross R: 752)

(17) Je peux toujours me trouver un ouvrage dans mon métier. (1971/94:281)

Encore plus spécifique est le sens d'un travail rémunéré accompli par un individu dont le poste porte un nom. Aux traits identifiés ci-dessus s'ajoutent ce nouveau trait, *n. Emploi* et *job* sont les deux mots employés dans se sens.

C. *t* ∩ *r* ∩ *n* [+travail, +rémunération, +poste portant un nom]: *emploi ~ job ~ (travail)*

(18) Puis cette année j'ai eu l'emploi de gérante. (Joanie O:325)

(19) "Bon on vient juste de mettre dehors notre secrétaire. Tu veux tu la job?". (Joanie O:331-2)

Quatrièmement, STB identifient un sens de travail qui combine les traits *t* et *r* non pas avec un poste nommé, mais avec la notion d'une tâche spécifique, *a. Travail* et *job* sont les noms les plus utilisés, avec un recours occasionnel à *emploi* et (pour les locuteurs natifs) *ouvrage*.

D. *t* ∩ *r* ∩ *a* [+travail, +rémunération, +tâche]: *travail ~ job ~ (emploi ~ ouvrage)*

(20) C'était un travail d'ordinateur alors c'était en anglais. (Kathy K:150-2)

(21) Il m'expliquait tout' les affaires, la compagnie, la job que je vas faire, comment je dois m'habiller, quelle heure je dois arriver. (Vincent W:692)

(22) . . . n'ont pas besoin de parler français du tout maintenant pour leur emploi. (Alicia M:855)

(23) D'abord que l'employé fait son ouvrage, il lui en demande pas plus. (1971/6:606)

En dernier lieu, il y a l'usage des mots de ce domaine pour référer au lieu du travail. *Travail* est le nom normalement utilisé ici, les locuteurs natifs employant aussi *job* et *ouvrage*.

E. $t \cap r \cap l$ [+travail, +rémunération, +lieu]: travail ~ (ouvrage ~ job)

(24) Je parlais français seulement à l'école quand j'étais petit mais maintenant c'est un peu à travail mais pas beaucoup. (Lynne B. 5-6)

(25) Il est venu me reconduire à l'ouvrage, puis il s'en allait chercher son patron. (1971/7:628)

(26) J'étais juste chauffeur pour s'en aller sur la job. (1971/32:171)

### 4.2 L'usage des anglophones et des francophones

Pour étudier ce champ sémantique, BST ont repéré les occurrences des mots chez les 120 locuteurs. Regroupant les sens B, C et D ci-dessus, c'est à dire, [travail rémunéré, ±nom, ±tâche], ils ont trouvé la distribution qui figure au Tableau 3. On y remarque une distribution plus équitable pour les différents mots que ce qui caractérisait l'emploi des verbes de résidence. *Travail*, utilisé un tiers du temps, est suivi par un emprunt de l'anglais, emprunt auquel on attribue le genre féminin en français québécois: *la job*. Les autres mots suivent avec moins de 15% des occurrences totales.

| Nom | Nombre de cas | Pourcentage |
|---|---|---|
| travail | 136 | 35% |
| job | 114 | 29% |
| ouvrage | 57 | 14% |
| emploi | 56 | 14% |
| poste/position | 32 | 8% |
| Total | 395 | 100% |

Tableau 3. Distribution des noms véhiculant les sens B, C et D [travail rémunéré, ±nom, ±tâche] chez les 120 locuteurs du corpus Sankoff-Cedergren 1971 (STB, pp.32-33).

Si les anglophones suivent le modèle courant dans la communauté francophone montréalaise, ils peuvent en quelque sorte bénéficier d'un mot de leur propre langue. Cependant, on pourrait imaginer qu'ils ne sont pas au courant de l'existence de cet emprunt si souvent utilisé par les francophones, ou qu'ils cherchent à éviter l'emploi d'un terme anglais en français. S'ils font une équivalence

entre leur *job* anglais et un mot français, ils choisiront peut-être le mot *emploi*. Encore une fois, si nous nous fions à la littérature, nous nous attendrons à ce que les anglophones aient un répertoire moins varié que celui des locuteurs natifs.

Le Tableau 4, qui présente les résultats, met en évidence une situation assez compliquée. (Les chiffres du Tableau 3 sont repris ici pour rendre plus facile la comparaison avec le groupe francophone.) Le fait que le premier choix du groupe anglophone (*travail*) soit le même que celui des francophones masque une différence importante entre les modèles d'usage des deux groupes. D'abord, nous remarquons que les anglophones ont tendance à éviter le mot *job*, ne l'employant qu'à un taux de 14%: la moitié du taux des francophones. Ils semblent le remplacer par *emploi*, utilisé à 34% (en comparaison avec le 14% des francophones). Deuxièmement, les anglophones évitent tout à fait un mot utilisé à 14% par les francophones: *ouvrage*. L'explication de cet évitement sera explorée ci-dessous. Notons en dernier lieu à propos du Tableau 4 que Vincent et Jocelyne semblent y afficher un usage plus proche des anglophones que des francophones. Ils n'ont aucune occurrence d'*ouvrage*, et ils affichent un taux d'usage du mot *emploi* qui dépasse de loin le taux des deux autres groupes. À mon avis, ces résultats s'expliquent au moins en partie par le nombre relativement faible d'occurrences globales dans ce domaine sémantique: Vincent et Jocelyne ensemble n'ont que 26 occurrences en tout. Le hasard veut qu'à l'époque de l'interview, Vincent "cherchait un emploi", et en parlait beaucoup, ce qui a certainement eu pour effet de gonfler son taux d'utilisation de ce mot.[6]

| Groupe de locuteurs | Mots | | | | | N |
|---|---|---|---|---|---|---|
| | job % | travail % | ouvrage % | emploi % | poste % | |
| 27 anglophones | 14 | 40 | 0 | 34 | 12 | 80 |
| Vincent/ Jocelyne | 19 | 8 | 0 | 58 | 15 | 26 |
| 120 francophones | 29 | 35 | 14 | 14 | 8 | 395 |

Tableau 4. Pourcentage des noms véhiculant les sens B, C et D [travail rémunéré, ±nom, ±tâche] chez trois groupes de locuteurs.

[6] Il est certes possible d'utiliser des mots autres qu'*emploi* dans cette expression, mais la plupart des locuteurs y affichent une préférence pour *emploi*.

L'interprétation de l'absence d'occurrences du mot *ouvrage* peut donner lieu à plusieurs explications. D'abord, il est possible que ce mot soit désuet à Montréal, résultat d'un changement linguistique qui aurait eu lieu pendant entre 1971 et 1993. Une recherche dans le corpus de 1984[7] a rapidement mis fin à ces spéculations; même si le travail quantitatif de comparaison des fréquences n'y est pas fait, le mot *ouvrage* y paraît maintes fois chez plusieurs locuteurs. Une deuxième voie d'explication serait la différenciation sociale dans l'usage de ce mot dans la communauté cible. Le Tableau 5 démontre qu'*ouvrage* est de loin préféré par les ouvriers (qui sont également ceux qui manifestent la plus haute fréquence de *job*).

| Sens | Mot | Locuteurs | | |
|---|---|---|---|---|
| | | professionnels | cols blancs, petits-bourgeois | ouvriers, chômeurs |
| travail rémunéré | ouvrage | 17% | 4% | 30% |
| | travail | 20% | 43% | 16% |
| | emploi | 39% | 23% | 11% |
| | job | 24% | 30% | 43% |
| tâche | ouvrage | 34% | 48% | 78% |
| | travail | 59% | 43% | 13% |
| | emploi | 3% | 0% | 0% |
| | job | 3% | 9% | 9% |

Tableau 5. Distribution sociale des mots portant le sens 'travail rémunéré' dans le corpus Sankoff-Cedergren de 1971 (STB, p.36).

Même si dans la communauté des locuteurs natifs, *ouvrage* s'emploie chez tout le monde, sa fréquence est nettement diminuée dans les classes les plus aisées, et plus particulièrement chez les cols blancs et petits-bourgeois, milieux où se trouvent la plupart de notre échantillon d'anglophones. Moins susceptibles de l'entendre de leurs pairs francophones, les anglophones n'auraient probablement pas rencontré non plus le mot *ouvrage* dans leurs cours de français. Faisant partie du vernaculaire de la classe ouvrière, les mots *ouvrage*

[7] Je tiens à remercier Hélène Blondeau de m'avoir aidée en dépouillant le corpus de 1984 pour rechercher les nombreux cas du mot *ouvrage*.

et *job*[8] sont alors moins disponibles aux anglophones que les mots plus standards.

## 5. Conclusion

Le lexique, nous l'avons vu, s'avère être un champ très propice à la mise en relation de deux langues dans une situation de contact linguistique. L'existence d'une étude quantitative détaillée des modèles qui sous-tendent les choix lexicaux chez les francophones (STB), et surtout la méthodologie développée dans cette étude, nous ont permis de faire une comparaison point par point de l'usage des deux communautés.

Les résultats de cette enquête sur l'usage du vocabulaire dans deux champs sémantiques nous ont permis de raffiner notre compréhension de ce qui caractérise le vocabulaire des locuteurs de L2. Dans le cas des verbes de logement, les anglophones possèdent le même éventail que les locuteurs natifs, mais manifestent un schème de préférences différent. Ils privilégient *habiter*, un mot qui a une équivalence en l'anglais *live*, et ne manifestent qu'un usage infréquent du mot préféré par les francophones: *rester*. En ce qui concerne les noms référant au travail, le répertoire des anglophones ne comprend que quatre des cinq mots utilisés par les francophones, puisque le mot *ouvrage* n'y figure pas. Cependant, cette population qui éprouve en général des sentiments positifs à l'égard de Montréal et aussi envers le français québécois (Thibault & Sankoff 1993) a réussi au moins partiellement à s'intégrer sur le plan linguistique. Le fait que les locuteurs de notre échantillon utilisent, quoiqu'à une faible fréquence, le nom *job* au féminin, et le verbe *rester* pour parler de la résidence, témoigne de leurs contacts avec la communauté majoritaire. Le français des anglophones est à mi-chemin entre le français qu'ils ont appris à l'école, et le français de la communauté francophone québécoise.

## Bibliographie

Archambault, Danielle, Catherine Faucher et Blagovesta Maneva. 1996. Fluency and use of segmental dialect features in the acquisition of a second language (French) by English speakers.

[8] Dans le cas de *job*, la "disponibilité" veut dire non pas la connaissance du mot en soi, mais plutôt la connaissance du fait que les francophones s'en servent couramment.

Actes du International Conference on Speech and Language Processing, Philadelphie, 4 octobre 1996.

Lambert, Wallace, J. Havelka & R. C. Gardner. 1959. Linguistic manifestations of bilingualism. *American Journal of Psychology* 72:77-82.

Sankoff, David, Gillian Sankoff, Suzanne Laberge, & Marjorie Topham. 1976. Méthodes d'échantillonnage et utilisation de l'ordinateur dans l'étude de la variation grammaticale. *Cahiers de linguistique de l'Université du Québec* 6:85-125.

Sankoff, David, Pierrette Thibault et Hélène Bérubé. 1978. Semantic field variability. In D. Sankoff, (réd)., *Linguistic Variation: Models and Methods*. New York: Academic Press, pp. 23-43.

Sankoff, Gillian, Pierrette Thibault, Naomy Nagy, Hélène Blondeau, Marie-Odile Fonollosa, et Lucie Gagnon. 1997. Variation in the use of discourse markers in a language contact situation. *Language Variation and Change* 9 (sous presse).

Thibault, Pierrette et Gillian Sankoff. 1993. Diverses facettes de l'insécurité linguistique: Vers une analyse comparative des attitudes et du français parlé par des Franco- et des Anglos-montréalais". *Cahiers de l'Institut de Linguistique de Louvain* 19,3-4:209-218.

Thibault, Pierrette et Diane Vincent. 1990. *Un corpus de français parlé*. Québec: Département de langues et linguistique, Université Laval.

Weinreich, Uriel. 1953. *Languages in Contact*. The Hague: Mouton.

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 147-160.

# La notion de « francisme » a-t-elle sa raison d'être?

*Sylvie Thiboutot et Pierre Martel*
Université de Sherbrooke

En novembre 1994, les marques lexicographiques ont fait l'objet d'une table ronde au cours de laquelle la question des marques topolectales fut longuement discutée. Par rapport aux positions diamétralement opposées de plusieurs linguistes avant cette rencontre, les discussions ont permis aux participants de faire des progrès convergents. Nous désirons poursuivre l'examen des marques topolectales en approfondissant la notion de « francisme ».

L'auteur du premier dictionnaire général du français québécois, Louis-Alexandre Bélisle, a marqué les mots spécifiques au Québec, ceux que l'on appelait alors canadianismes. Sauf le *Petit Breton*, qui repérait tant les canadianismes que les francismes, les dictionnaires subséquents, soit le *CEC Jeunesse*, le *CEC Intermédiaire*, le *Dictionnaire du français Plus* (*DFP*) et le *Dictionnaire québécois d'aujourd'hui* (*DQA*) ont adopté une autre politique éditoriale : dans ces ouvrages, les québécismes, comme il convient maintenant de les appeler, ne sont plus marqués; seuls les francismes le sont. Cette position nous semble discutable et ambiguë; c'est pourquoi nous proposons une autre façon d'indiquer la variation géographique : le marquage des mots, expressions et sens québécois. Loin d'être un retour en arrière, cette façon de faire constitue une vision renouvelée du français québécois.

Pour Jean-Claude Boulanger (1994), marquer les québécismes équivaudrait à une marginalisation de ces mots et au renforcement du sentiment d'insécurité linguistique des Québécois et Québécoises. Ce type de marquage signifierait en outre qu'il existe une norme supérieure centralisée ailleurs, à Paris. Pour ce lexicographe, mettre en évidence les unités linguistiques québécoises reviendrait à les rendre irrecevables, car les mots marqués sont perçus négativement par les usagers des dictionnaires : un mot identifié comme québécisme devient automatiquement un mot à proscrire. Pour ces raisons, les auteurs du *DQA*, à l'instar de ceux du *DFP*, ont choisi de marquer les mots qu'ils

pensaient être en usage en France seulement.

Depuis la parution du *DQA*, des linguistes ont mis en doute le bien-fondé de marquer les francismes, entre autres Claude Verreault (1994) et Annette Paquot (1995). Ceux-ci ont démontré que le marquage à l'aide des marques France ou Surtout en France est imprécis, notamment en ce qui concerne le choix des mots à marquer. Annette Paquot a comparé, pour la lettre **M**, les mots marqués France dans le *DQA* (120 mots ou emplois) et dans le *DFP* (30 mots ou emplois) : les deux ouvrages ne sont d'accord que pour 15 mots. Claude Verreault, quant à lui, constate, dans le *DQA*, l'absence de marques à des mots qui semblent inconnus au Québec (**jus de chaussette** « mauvais café », **pavé** « gros livre indigeste ») et le marquage des mots qui ne sont pas (ou plus) réservés à la France.

À ces voix, il faut ajouter celles des lecteurs et lectrices de ces derniers dictionnaires québécois, notamment les critiques du *DFP* et du *DQA*, qui voudraient savoir si tel mot ou tel sens est limité au Québec ou s'il est panfrancophone.

Lionel Boisvert écrivait, à la suite de la parution du *DFP* :

> *Bazou* et *guimbarde* sont donnés tous deux comme familiers, mais rien ne nous indique que l'un des deux est d'usage strictement québécois, contrairement à l'autre; rien ne nous dit que *maganer* ne s'emploie qu'au Québec, mais que *se poiler* y est inconnu. [...] L'absence quasi généralisée de telles indications nous place en porte-à-faux entre deux usages lexicaux. (Boisvert 1988 : 29)

À la table ronde sur les marques lexicographiques, plusieurs intervenants ont souligné l'importance de marquer les particularités québécoises, comme en font foi les extraits suivants :

> Le fait de signaler qu'un mot n'est pas employé par l'ensemble de la communauté francophone ne jette pas

> l'opprobre sur ce mot, la marque (Québec) ou (Canada) ne constitue pas une marque d'infamie, c'est tout simplement une information qui est communiquée à l'usager du dictionnaire. (De Villers 1994 : 5)

> L'étudiant a besoin de savoir si le mot qu'il utilise ou qu'il lit appartient au fond commun international actif ou est propre au Québec, ou au français de quelque autre région de la francophonie [...] Le but est clair : il ne faut pas marginaliser le jeune Québécois, mais il ne faut pas non plus l'isoler ou le berner en lui faisant croire par exemple que PATENTE a le sens de TRUC dans toute la francophonie (Gadbois 1994 : 6)

À cette même table ronde, Hélène Cajolet-Laganière (1994) faisait part des résultats d'une étude faite auprès de 110 personnes, qui portait sur les attentes et besoins du public québécois en matière de dictionnaires de langue. À la onzième question, on proposait aux répondants trois choix de dictionnaires : 85 % d'entre eux désirent un dictionnaire général dans lequel les mots et les sens employés au Québec seraient identifiés à l'aide d'une marque quelconque qui les distingueraient des mots et sens employés dans les autres pays francophones.

Selon nous, il est difficile de maintenir la notion de « francisme » comme politique éditoriale pour plusieurs raisons. Tout d'abord parce que nous sommes francophones à part entière et qu'il semble difficile d'« exclure » de notre lexique potentiel ou virtuel des mots ou des sens, sous prétexte qu'ils ne sont pas (ou peu) employés en français québécois.[1] C'est là un point important : nous croyons que tout « francisme » peut être attesté dans un texte en français québécois, que

[1] En effet, En France, signifie pour le lecteur que le mot ou le sens appartient à une autre variété de français et qu'en conséquence, il ne fait pas partie de « notre » langue.

l'on pense aux romans de Jacques Godbout, où il utilise **W.C.** ou à ceux de Michel Tremblay où figure **congère**.

Tout Québécois et Québécoise entend des « discours » ou lit des textes en provenance de France : les dictionnaires, les textes littéraires, les revues, les magazines, les films, les interviews, etc. Tout ce vocabulaire du français de France existe au moins sous la forme passive chez les Québécois et Québécoises. Or le vocabulaire passif fait partie du lexique individuel (au même titre que les vocabulaires actif, disponible et virtuel). De plus, l'examen de vastes corpus révèle fréquemment que des unités de fréquence 0 (mots absents) passent à la fréquence 1 (hapax), puis à la fréquence 2, 3, 4...

En outre, quand un lexicographe québécois indique France, ou Surtout en France, il commet souvent une faute par restriction, car pour la plupart de ces termes, l'extension géographique s'étend aussi à la Belgique, à la Suisse et parfois à l'Afrique.

Maintenir une marque dont l'application est à ce point approximative enlève une partie de la crédibilité à l'ouvrage dictionnairique. La notion de francisme suppose de plus une connaissance de l'usage du français en France, ce que nous, Québécois et Québécoises, n'avons que partiellement.

Pour vérifier si effectivement les mots que les dictionnaires qualifient de francismes peuvent être diffusés au Québec, nous avons recherché les 1270 « francismes » du *DQA* dans un corpus composé des textes de cinq journaux et magazines québécois (CD Actualité/Québec) pour l'année 1992 (Thiboutot 1995).

Certains mots cherchés sont absents, mais la plupart sont bel et bien attestés. **Week-end**, sans marque dans le *DFP*, mais marqué Surtout en France dans le *DQA* est un exemple des plus intéressants : en 1992, ce mot apparaît dans 2667 articles, et en 1993, pour un corpus équivalent, dans 3552. Il est certain que la fréquence de ce mot va en augmentant au Québec. Des mots marqués « francismes » comme **hectare**, **élections législatives**, **petit déjeuner**, **fisc**, **pub**, **autocar** figurent dans plus de 200 documents. Par contre, d'autres mots sont peu attestés : **contre-filet** (2 documents) **friperie** (7 documents), **grog** (1 document), mais leur fréquence est non nulle. Même les mots qui

sont encore absents du corpus québécois sont susceptibles d'y apparaître lorsqu'on augmentera sa taille : **barbet,** tant dans son sens non marqué « espèce d'épagneul » que dans celui « variété de rouget » marqué France est absent du corpus. De même, **barbe à papa**, « confiserie » est absent du corpus étudié, mais, selon nous, est connu et utilisé au Québec.

Nous contestons plus particulièrement l'appellation de francisme ou la marque France parce que le fait qu'un mot se rapporte à une réalité typiquement hexagonale ne signifie pas pour autant que ce mot est perçu comme étant étranger par la plupart des Québécois et Québécoises, contrairement aux choses (aux référents) qui sont exclusives à la France. Les mots **mistral** et **minitel** sont le bien de tous les francophones, tout comme **girafe, baobab, kangourou**, etc. qui appartiennent à tous les francophones, et donc aux Québécois et Québécoises, même si les référents sont typiques à des aires limitées. La fréquence relativement élevée de mots comme **agent de police**, **chambre des députés**, **élections législatives**, **gendarme**, **maréchal**, etc. dans le corpus journalistique étudié montre que ces mots sont répandus, et par conséquence, connus au Québec. En parcourant le corpus, nous sommes rendu compte que des réalités que l'on peut croire à première vue réservées à la France peuvent être empruntées, « importées », comme c'est le cas pour **lycée**, **arrondissement**, **chef-lieu**, **préfecture**, **T.V.A.** En effet, il y a, au Québec, deux lycées; Montréal est divisée, sur le plan administratif, en 9 arrondissements, avec la création des M.R.C. et du poste de préfet de comté sont nées deux nouvelles appellations, en vigueur dans certaines parties du Québec : **chef-lieu** et **préfecture**; on a lu dans les journaux que la **T.V.A.** « taxe à la valeur ajoutée » pourrait remplacer la TPS. On pourrait également parler de produits « importés » de France, comme le **bateau-mouche**, qui est devenu une réalité québécoise, ou comme le **TGV** « train à grande vitesse » qui pourrait le devenir.

Pour toutes ces raisons, nous remettons en question la pertinence de la marque *francisme*, même si, a priori, il pouvait sembler plus simple d'identifier les francismes plutôt que les québécismes.

De plus, l'argument selon lequel un mot marqué **Q** serait nécessairement perçu comme condamnable ne sera plus valable le jour où sera décrit le français québécois standard, c'est-à-dire la norme du français québécois (Martel et Cajolet-Laganière 1995).

Il y a lieu de se demander aussi quels sont ces mots qui sont considérés comme des francismes? Bien qu'ils soient parfois sentis par certains Québécois et Québécoises comme appartenant au français de France, il apparaissent néanmoins dans des textes québécois, mais à une très faible fréquence. Ces mots, parfois rares, sont utilisés surtout dans des textes soutenus, littéraires, spécialisés ou techniques par une minorité de Québécois et de Québécoises; ils sont en outre connus, sous forme passive, par un plus grand nombre d'entre eux.

Dans un projet de dictionnaire français au Québec, il faut hiérarchiser tous les usages de cette variété nationale autour de l'usage légitime québécois comme critère d'acceptabilité d'un mot ou d'un sens. Serait-ce à dire qu'il ne faut pas marquer ce qui est proprement québécois? Nous ne le croyons pas, parce que, pour certains mots ou certains sens, les Québécois et Québécoises ressentent le besoin de se situer par rapport au français de référence ou auraient avantage à connaître l'écart entre notre usage et celui de France, à une époque où les échanges sont plus intenses entre ces deux communautés, pour au moins pouvoir s'adapter aux exigences de la situation de communication. (Corbeil 1994 : 16)

Nous proposons des solutions multiples de marquage, selon les cas, mais basées sur les principes suivants.

Il faut marquer par rapport aux occurrences des mots dans les discours/textes québécois. La plupart des mots repérés comme des francismes appartiennent à la même catégorie que tous les autres en situation de concurrence linguistique ou de synonymie (par exemple : char, voiture, automobile, rencontrés dans un corpus québécois).

Il faut maintenir la marque **Q** et l'indiquer en fonction du français référentiel et non par rapport aux français de France, de Belgique, etc. que l'on ne connaît pas ou que l'on connaît mal. (Martel et Cajolet-Laganière 1996). On ne peut cependant pas marquer tout ce qui est « québécois ». Il est impensable de marquer des mots comme

**hiver** ou **fleuve**, des mots dont les différences sont de nature connotative et encyclopédique.

Les mots marqués **Q** seront ceux dont la différence est importante (signifiant, signifié, niveau de langue, etc.), dont l'écart se révèle important dans les communications orales ou écrites. Un mot comme **inhalothérapeute** serait marqué, car l'extension de la forme est limitée au Québec, même si le mot est standard. L'aire d'extension d'un mot, d'un sens, est une réalité linguistique à décrire, au même titre que le genre, l'orthographe, etc. L'indication de l'origine québécoise du mot pourrait parfois figurer dans la marque étymologique.

Le fait de décrire le français québécois par rapport au français de référence et non au français de France contribuerait à lever l'ambiguïté pour certains mots ou sens qui ne sont pas propres au français québécois. En effet, nous savons, mais pas toujours, que certains mots ont déjà appartenu au français de France ou à l'un ou l'autre de ses parlers (**jaser** au sens de parler), d'autres mots peuvent être actuellement en usage en France ou ailleurs (**abri d'auto** dans un sens proche de celui du français québécois serait attesté dans le sud de la France). Bien entendu, il faut prendre garde de tomber dans le piège de certains dictionnaires précédents, comme celui de Bélisle, qui identifiaient les québécismes, le plus souvent pour les dévaloriser, même si ce n'était pas là leur intention.

Sauf quelques rares exceptions, les mots ne comporteront pas plus de 2 marques. Certaines de ces marques, jusqu'à maintenant indiquées comme topolectales, deviendraient sociolectales, même si cette catégorie de marques est difficile à utiliser : il faut éviter à tout prix d'identifier les mots comme étant québécois et donc familiers, par exemple de marquer **mitaine** (familier) et **moufle** (standard). On devrait plutôt chercher le type de textes, le type de communication, l'appartenance d'un mot à la langue de spécialité ou la fréquence d'emploi dans les textes du français québécois. Prenons par exemple **perdrix** et **gélinotte huppée**, le premier est le mot courant, utilisé dans la langue de tous les jours, tandis que le deuxième est un terme plus officiel, plus spécialisé, que l'on peut trouver, par exemple, dans les brochures du Ministère. Citons aussi les noms des repas : nos

**déjeuner, dîner et souper** deviennent, dans des textes officiels, des **petit déjeuner, déjeuner et dîner**.

Dans certains cas, la référence géographique pourrait apparaître à l'intérieur des définitions (même si, à certains égards, ce trait géographique ne peut être considéré comme une marque), que l'on décrive une réalité de France ou du Québec. Ainsi à **lycée**, on aurait « Dans le système d'enseignement de France », et à **cégep** : « Dans le système d'enseignement du Québec », **cassoulet** « ragoût typique de la cuisine française » **poutine** « ... de la cuisine québécoise ». On éviterait ainsi le redoublement de marques, comme c'est le cas, dans le *DQA*, du mot **sirocco**, dont la définition est précédée de (Afrique), alors que le sens comporte la dimension géographique « d'origine saharienne ». Il faudrait donc distinguer la marque topolectale réservée au signe linguistique du référent géographique à inclure dans la définition, s'il y a lieu. De même, il faudrait faire réellement la distinction entre les référents français (**minitel**, **lycée**) et les signes linguistiques utilisés surtout en France (**moufle**, **tennis**).

D'autres marques du type **peu** (ou **moins**) **fréquent**, **récemment emprunté au français de France** pourraient aussi être utilisées.

D'autres solutions de marquage nous paraissent utiles. Par exemple, dans le *DQA*, le traitement du mot **palette** est des plus intéressants. La première définition est précédée de (coupe nord-américaine) « pièce de boeuf [...] », tandis qu'avant la deuxième, on a (coupe française) « pièce de viande de mouton, de porc [...] ».

Nous proposons quelques exemples de ce que pourrait être ce nouveau système de marquage, en comparaison avec le système adopté par le *DQA*.

Dans la proposition qui suit, nous ne préjugeons pas de l'acceptabilité ou non des anglicismes, de la classification juste des sens, de l'exactitude de la définition, etc. Cet essai vise uniquement à établir un nouveau système de marques topolectales et, par ricochet, des marques sociolectales qui devraient accompagner la marque **Q**, quand cela est nécessaire.

| DQA | PROPOSITION |
|---|---|
| ***ANGINE*** [...] n. f. **1.** (Cour. en France) Inflammation de la gorge. **2.** *Angine de poitrine*, douleurs dans la région du cœur, accompagnées d'angoisse. => **infarctus.** *Faire de l'angine. Crise d'angine.* | ***ANGINE*** [...] n. f. **1.** Q. Douleurs dans la région du cœur, accompagnées d'angoisse. => **infarctus.** « Tu sais ce que le docteur a dit quand ta mère a eu sa dernière crise d' **angine** qu'il fallait la ménager / tu veux pas être responsable de la mort de ta mère... », J. Bertrand, *Mon Tarzan, toi Jane...: 75, Faire de l'angine. Crise d'angine.* ANGINE DE POITRINE, même sens. **2.** Peu f, inflammation de la gorge. « Je la retrouve dans une petite salle du Concorde, grelottant, tenant un chandail enroulé autour de son cou frêle. **Angine**..., souffle-t-elle » *Le Soleil*, 18/07/1992. => **amygdalite** ° |
| ***ARRONDISSEMENT*** [...] n. m. ... **2.** (France) Circonscription administrative. *Le département français est divisé en un certain nombre d'arrondissements. Chef-lieu d'arrondissement.* => **sous-préfecture.** — Subdivision administrative dans certaines grandes villes (Paris, Lyon, Marseille). *Le V^e^, le XVI^e^ arrondissement à Paris.* | ***ARRONDISSEMENT*** [...] n. m. ... **2.** Circonscription administrative dans certaines grandes villes, à Paris, Lyon, Montréal... => **quartier.** « Balcons superposés et escaliers en façade: voilà le menu régulier que propose, chaque mois, l'ordre du jour des assemblées des comités-conseils qui siègent dans les neuf **arrondissements** montréalais. », *La Presse*, Édit. 11/10/1990, A. Gruda. |
| ② ***BASKET*** n.f. [...] ■ (France) Anglic. Chaussure de sport en toile moulant la cheville, à semelle et rebords de caoutchouc. => **espadrille, tennis.** —Loc. fam. *Être à l'aise dans ses baskets*, être décontracté. *Lâche-moi les baskets*, laisse-moi tranquille. | ② ***BASKET*** n.f. [...] Peu f. Syn **espadrille.** « La pédale au fond, le béret bien enfoncé, les **baskets** bien lacés, il roule sur le circuit du Metropolis, suscitant les ho! et les ha! à chaque accélération ... » *La Presse*, 10/12/1992. —Loc. fam (peu f) : *être à l'aise dans ses baskets*, être décontracté; *lâche-moi les baskets*, laisse-moi tranquille. |
| ***BATEAU*** [...] n. m. **1.** ... — (France) BATEAU-MOUCHE : bateau transportant des passagers sur la Seine à Paris. *Des bateaux-mouches.* ... | ***BATEAU*** [...] n. m. **1.** ... — BATEAU-MOUCHE : bateau transportant des passagers sur l'eau. *Des bateaux-mouches.* « ... Dans ce que l'on appelle l'arrière-pays de Charlevoix, l'une des excursions les plus spectaculaires demeure la croisière en **bateau-mouche** sur la rivière La Malbaie au coeur des Hautes-Gorges de Charlevoix. » *La Presse*, 11/07/1992. |

| DQA | PROPOSITION |
|---|---|
| ***CASSOULET*** [...] n. m. ■ (France) Ragoût préparé avec de la viande (confit d'oie, de canard, mouton ou porc) et des haricots blancs. | ***CASSOULET*** [...] n. m. Mets typique du sud de la France consistant en un ragoût de viande (confit d'oie, de canard, mouton ou porc) et de haricots blancs. ▫ La Râpière c'est, à Montréal, la maison des confits et des **cassoulets**. ▫ *La Presse*, 15/01/91. |
| ***ESPADRILLE*** [...] n f. **1.** Chaussure de sport souple (basse => **tennis** ou moulant la cheville => **basquet**), en toile, à semelle et rebords de caoutchouc *Une paire d'espadrilles de tennis*. — REM. Cet emploi est critiqué. **2.** (France) Chaussure dont l'empeigne est de toile et la semelle de corde. *Une paire d'espadrilles*. | ***ESPADRILLE*** n.f. [...] Q. Chaussure de sport souple ou moulant la cheville, à semelle et rebords de caoutchouc. ▫ Les chandails et les chemises coupés à la taille, les t-shirts et les camisoles, les jeans, **espadrilles** et pantalons moletonnés ne seront donc plus tolérés ▫ *La Presse*, 1 /11/1993. => **basket** peu f., **tennis** peu f |
| ***LYCÉE*** [...] n. m. (France) **1.** Établissement public d'enseignement secondaire (classique, moderne ou technique). Les professeurs d'un lycée. => **cégep, polyvalente, secondaire.** ▫ collège. **2.** Époque des études secondaires. *Elle ne l'a pas revu depuis le lycée*. ▫ **lycéen, enne** ... | ***LYCÉE*** [...] n. m. **1.** Établissement public d'enseignement secondaire (classique, moderne ou technique) dans le système français. ▫ L'Association des collèges privés du Québec représente 25 des 26 collèges privés, en plus de deux **lycées**, soit quelque 20 000 des 160 000 étudiants de niveau collégial au Québec ▫ *La Presse*, 3/11/1992. *Les lycées de Montréal décernent un diplôme français*. => **cégep, polyvalente, secondaire, collège**. |
| ***MINITEL*** [...] n. m. ■ (France) Petit terminal de consultation de banques de données commercialisé par les P.T.T. *Des minitels*. — REM. Ce mot est un nom de marque déposée. | ***MINITEL*** [...] n. m. Petit terminal de fabrication française qui sert à consulter des banques de données. ▫ L'Université de Montréal est devenue hier la première institution universitaire nord-américaine à faire son entrée sur le réseau télématique français **Minitel**, auquel sont reliés près de sept millions de foyer ▫ *Le Devoir*, 29/01/92. *Des minitels*. — REM. Ce mot est un nom de marque déposée. |

| DQA | PROPOSITION |
|---|---|
| ① ***SOUPER*** [...] v. intr. ▬ conjug. 1. **1.** Prendre le repas du soir. ① => **dîner** (2). *Qu'est-ce qu'on mange pour souper? Souper au restaurant.* **2.** (France) Prendre un repas ou une collation à une heure avancée de la nuit, après un spectacle, au cours de la soirée.=> anglic. **luncher.** *Les comédiens vont souper dans un restaurant chic.* **3.** Fig. Loc. fam. En avoir soupé (de qqch.), en avoir assez, en être excédé (→ Avoir son voyage; en avoir marre) ▪ ② **souper** n. m. **1.** Repas du soir. => ② **dîner.** *Un souper d'anniversaire. Elle téléphone toujours à l'heure du souper. — Des soupers-bénéfices*.* **2.** (France) Repas, collation pris(e) tard en soirée. => anglic. **lunch.** *Un souper aux chandelles.* **3.** Les mets qui composent le souper (1, 2). *Faire réchauffer le souper. Mettre son souper au micro-ondes.* — REM. Dans les circonstances officielles, on emploie *dîner* (2) au lieu de *souper* (1). | ① ***SOUPER*** [...] v. intr. Conjug. 1. **1.** Q (f dans la francophonie). Prendre le repas du soir. **2.** Peu f. Prendre un repas ou une collation à une heure avancée de la nuit, après un spectacle, au cours de la soirée. **3.** Fig. Loc. fam. En avoir soupé (de qqch.), en avoir assez, en être excédé (→ Avoir son voyage; en avoir marre) |
| | ② ***SOUPER*** n.m. **1.** Q (f dans la francophonie). Repas du soir. *Un souper d'anniversaire.* ▪ Craignant de revenir tard, elle demanda à Rachel et à Bohu de descendre chez elle vers cinq heures pour accueillir Denis à son retour de l'école et partager avec lui le souper qu'elle avait mis sur le feu. ▪ Y. Beauchemin *Juliette Pomerleau* : 656. => **dîner**. – Des soupers-bénéfices*. **2.** Peu f. Repas, collation pris(e) tard en soirée. **3** Les mets qui composent le souper (1, 2). *Faire réchauffer le souper. Mettre son souper au micro-ondes.* |
| ***WEEK-END*** [...] n. m. ■ Anglic. (Surtout en France) Congé de fin de semaine, comprenant le samedi et le dimanche. *Nous partons à la campagne tous les week-ends.* | ***WEEK-END*** [...] n. m. Syn. **fin de semaine*** de plus en plus f. ▪ Et son congrès de Sherbrooke, ce **week-end**, pourrait être l'occasion d'amorcer cette nécessaire transformation en profondeur ▪ *La Presse*, Édit. 24/05/1991, P. Gravel. |



157

Pour un mot comme **lycée,** dont le référent se situe en France, nous avons utilisé le marquage intradéfinitionnel. De même, pour **minitel** et **cassoulet**, nous avons ajouté les traits d'origine ou d'identification à l'intérieur de la définition.

Pour certains autres mots, comme **angine**, **souper**, **espadrille**, **basket,** nous avons indiqué le caractère de « rareté » par des marques de fréquence. Par exemple, pour **angine**, nous avons interverti l'ordre des sens, et fait précéder le sens 1 « douleurs dans la région du coeur [...] » de la marque **Q**, et le deuxième sens « inflammation de la gorge » de peu fréquent. Pour le substantif **souper**, le sens 1 « repas du soir » est marqué Q (fréquent dans la francophonie) alors que le sens 2 « repas, collation pris (se) tard en soirée » est marqué peu fréquent. Dans de rares exceptions, nous ferons référence au reste de la francophonie; par exemple, pour le mot **souper**, marqué vieux, régional (Belgique, Canada, Suisse, etc.) dans le *Nouveau Petit Robert*, qui est plus qu'un québécisme.

Nous avons par ailleurs supprimé la marque topolectale lorsque des mots ne sont plus réservés à l'usage de France, et nous avons ajouté des extraits de textes québécois pour montrer que ces réalités sont maintenant nôtres, comme **arrondissement**, bateau-**mouche**. Pour **week-end**, nous avons indiqué que le synonyme **fin de semaine** est plus fréquent, ce qui laisse entendre que **week-end** est tout de même utilisé au Québec.

## CONCLUSION

Même si les unités lexicales marquées *France* ou *Surtout en France* dans le DQA peuvent être senties par plusieurs Québécois et Québécoises comme typiques au français de France, nous ne croyons pas que nous devons les décrire en fonction du corpus français, qui n'est pas notre objet de description et que nous connaissons mal.

Nous constatons, au contraire, que ces unités appartiennent au français québécois, car la plupart d'entre elles y sont attestées ou le deviendront en augmentant la taille et la représentativité de la Banque de données textuelles de Sherbrooke (BDTS). Nous avons donné

aussi d'autres arguments qui appuient notre parti pris d'une description interne et complète du français québécois.

Une fois le principe accepté, reste la manière de marquer ces mots et ces sens. Nous croyons que diverses stratégies de marquage doivent remplacer l'ancienne marque France. Nous en avons proposé quelques-unes : la marque intra-définitionnelle, l'ajout des traits d'origine ou d'identification à la France, l'utilisation des marques de fréquence ou de spécialisation.

Cette façon de marquer est une hypothèse, une nouvelle voie que nous ouvrons. Notre réflexion continue et nous attendons des réactions à ces propos.

## Bibliographie

Boisvert, Lionel. 1988. « Le français, plus... », dans *Circuit*, décembre, p. 29-30.

Boulanger, Jean-Claude. 1994. « Le pacte normatif du français québécois : réflexions sur les marques lexicographiques diatopiques », Communication présentée lors de la *Table ronde sur les marques lexicographiques*, Montréal, novembre, (à paraître).

Cajolet-Laganière, Hélène. 1994. « Attentes et besoins du public québécois en matière de langue ». Communication présentée lors de la *Table ronde sur les marques lexicographiques*, Montréal, novembre, (à paraître).

Corbeil, Jean-Claude. 1994. « Les marques d'usage comme technique de description des aspects connotatifs du lexique », Communication présentée lors de la *Table ronde sur les marques lexicographiques*, Montréal, novembre, (à paraître).

Gadbois, Vital. 1994. « Les marques d'usage en lexicographie pour les étudiants et les étudiantes du collégial québécois ». Communication présentée lors de la *Table ronde sur les marques lexicographiques*, Montréal, novembre, (à paraître).

Martel, Pierre et Hélène Cajolet-Laganière. 1995. « Oui... au français québécois standard », dans *Interface*, Automne, p. 14-25.

Martel, Pierre et Hélène Cajolet-Laganière. 1996. *Le français québécois. Usages, standard et aménagement*, Québec, Institut québécois de recherche sur la culture, Diagnostic, n° 22, 141 p.

Paquot, Annette. 1995. « Le terme *francisme* : les avatars lexicographiques d'un concept à revisiter », dans *Langues et linguistique*, n° 21 p. 123-136.

Thiboutot, Sylvie. 1995. *La pénétration du français de France dans les textes journalistiques québécois*. Mémoire de maîtrise, Université de Sherbrooke, 115 p.

Verreault, Claude. 1994. « Inclusion, reconnaissance et identification des francismes dans les dictionnaires québécois : problèmes et méthodes à la lumière de l'expérience du *Dictionnaire québécois d'aujourd'hui* », Communication présentée lors *4e colloque international « Français du Canada - français de France. Lexicologie canadienne : d'hier à aujourd'hui »*, Chicoutimi, Université du Québec à Chicoutimi, septembre.

Villers, Marie-Éva de. 1994. « Les marques lexicographiques : des points de repère essentiels pour l'usage des mots ». Communication présentée lors de la *Table ronde sur les marques lexicographiques*, Montréal, novembre, (à paraître).

# Section 4

*Emprunts*

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 163-176.

# L'absence de flexion sur les emprunts à l'anglais dans le français cadjin

*Sylvie Dubois et David Sankoff*
Louisiana State University et Université de Montréal

**1. Introduction.** Bien qu'elle ne porte que sur la prononciation, la citation ci-dessous reflète des stéréotypes plus généraux sur la non-intégration phonologique, morphologique et syntaxique d'emprunts à l'anglais dans le français cadjin parlé dans le sud de la Louisiane. L'objectif de cet article est d'analyser l'usage réel de l'emprunt en français cadjin dans le but de vérifier les processus d'intégration de mots d'origine anglaise.

> Au Canada, j'ai remarqué qu'il existe une plus grande tendance à franciser les anglicismes adoptés. On tâche de changer leur nationalité et de les revêtir à la française. En Louisiane, malheureusement, nous ne prenons pas même la peine de franciser les mots empruntés à l'anglais. Pour la plupart, nous les acceptons tels quels sans souci du danger que nous courons. Le canadien va tâcher de prononcer les mots en anglais un peu selon les règles de prononciation de la langue française. Le louisianais, lui, quand il parle français, intercale le mot anglais prononcé tout simplement comme il l'est en anglais. Outre l'accent français qui se fait entendre (surtout dans la lettre R), l'anglicisme dans le parler franco-louisianais est à peu près de l'anglais pur. Voilà encore, raison de plus, pourquoi je dis que l'anglicisme, dangereux au Canada, est réellement mortel en Louisiane. Le grand danger est non seulement le nombre illimité d'anglicismes qui entrent dans la langue mais aussi la manière dans laquelle ils y entrent – c'est-à-dire 'à l'anglaise' et non 'à la française'. (Pellerin 1937:150)

Notre discussion s'insérera dans les débats sur la distinction entre les alternances de langue *"code-switching"* et les emprunts (ex: Sankoff et al. 1990, Meechan & Poplack 1995, Poplack & Meechan 1995). Nous adoptons comme hypothèse de travail que les mots d'origine anglaise qui se trouvent seuls dans du discours français, surtout dans un contexte où leur fonctionnement syntaxique et sémantique dépend des mots grammaticaux français qui les entourent, sont des emprunts et non pas des alternances (Poplack et al. 1988). Par contre, les plus longs fragments du discours ou des

fragments de phrase qui ne contiennent que des mots grammaticaux (mots de fonction) et des mots lexicaux (mots de contenu) d'origine anglaise sont clairement des alternances. Nous exclurons ces derniers de notre étude et nous nous concentrerons sur les mots isolés. Nous évaluerons jusqu'à quel point ces mots satisfont aux critères traditionnels des emprunts: mots lexicaux plutôt que mots grammaticaux, acceptation ou non à l'intérieur de la communauté, intégration phonologique et morphologique au français. Notre méthodologie quantitative libérera notre discussion d'une dépendance sur un ou deux exemples et nous permettra de discerner des tendances significatives à l'intérieur de la communauté cadjine.

**2. Le corpus.** Cette étude préliminaire vise à planifier un plus vaste projet (Dubois 1995, Dubois et al. 1995a&b, Dubois et Melançon 1996) portant sur l'usage de l'emprunt et l'alternance de langues dans plusieurs variétés de français parlées en Louisiane telles que le français cadjin, le français créole et la variété de français parlée par les Amérindiens Houmas. Notre corpus pilot est constitué de cinq enregistrements en français cadjin effectués entre 1994 et 1995. Nous avons transcrit les conversations et séparé le discours majoritairement en anglais du discours en français. Ensuite nous avons compilé le nombre de mots dans chaque entrevue ainsi que le nombre et l'origine (anglaise/française) des emprunts. Le tableau 1 montre la distribution des mots et le nombre d'emprunts d'origine anglaise et française dans le corpus et dans chaque entrevue.

| Entrevues | Mots | Mots en discours français | | Emprunts à l'anglais | | Mots en discours anglais | | Emprunts au français | |
|---|---|---|---|---|---|---|---|---|---|
| | | N | % | N | % | N | % | N | % |
| GRATTON | 13933 | 13762 | 98.7 | 69 | 0.5 | 99 | 0.7 | 3 | 3.0 |
| BREAUX | 11272 | 10795 | 95.7 | 242 | 2.2 | 234 | 2.1 | 1 | 0.5 |
| RODRIGUE | 10302 | 7588 | 73.6 | 261 | 3.4 | 2442 | 23.7 | 11 | 0.5 |
| TEMPLET | 4876 | 3702 | 75.9 | 41 | 1.1 | 1131 | 23.2 | 2 | 0.2 |
| MOUNIER | 4051 | 4020 | 99.2 | 31 | 0.8 | 31 | 0.8 | 0 | 0.0 |
| TOTAL | 44434 | 39867 | 90.0 | 644 | 1.6 | 3937 | 0.8 | 17 | 0.4 |

Tableau 1. Nombre de mots et d'emprunts dans les sections français et anglais des entrevues.

Le taux global d'emprunts à l'anglais est tout à fait similaire à celui trouvé en français parlé à Montréal et dans la région Ottawa-Hull (Poplack et al. 1988). Il est fort intéressant de constater que les deux seules entrevues menées par des interviewers membres de la communauté cadjine, Gratton et Mounier, sont celles qui contiennent le moins de discours anglais et, encore plus important, ont les plus bas taux d'emprunts à l'anglais.

**3. Les mots empruntés.** Les mots empruntés à l'anglais appartiennent à plusieurs catégories grammaticales.

**• Les noms**

(1) Tous les **grandchildren** parlent en français (ROD:19/31)
(2) Elle est folle de ça, une vieille volaille pour faire de la fricassée plutôt que ces **fryers** qu'on achètent tu sais. (BRE:77/18)
(3) J'ai cuit dessus (dessur) mon **stove**, elle elle avait un **stove** à bois
(4) Elle travaille pour le beau-père dans un **office**. (BRE:122/21)
(5) On n'a que nos **kin** et nos tantes qui parlaient pas en anglais. (GRA:2/4)
(6) Est-ce c'est un bon **choice** que les cadiens se sentent assimilés. (GRA:27/19)
(7) Ils ont trouvé que c'était un **gland** cancer. (MOU: 14/90)

**• Les expressions nominales figées.** Puisqu'elles sont enchâssées dans un groupe nominal, nous les avons considérées comme des emprunts plutôt que des alternances:

(8) Tu goûtes que le **hot sauce**, tu goûtes que le tabasco (ROD:8/24)
(9) Tu vois le **Pledge of Allegiance**, moi je le sais pas dans le français (ROD:21/31)
(10) Ils ont suit le **King of England**, le roi. (GRA:11/33)
(11) C'était des **gravel roads** qu'on allait: en allant à Morgan city (BRE:32/15)
(12) Quand mon père a été travaillé sur le **dredge boat**, il faisait encore pas un tas d'argent (ROD:49/35)
(13) Les vieux **air-conditioners** aussi, les vieux **washing machines** avec le **washboard** (BRE:50/16)
(14) T'achetais pas de **cooking oil**. Pas de **vegetable oil**. (BRE:67/17)
(15) On met des tomates et nos oignons et tout ça et on fait un **brown gravy**. (BRE:91/19)
(16) Elle a appelé un **French teacher**. (BRE:116/21)

(17) Il y a des **programs** sur **Channel Twelve** des fois en français .(BRE:6/9)
(18) Il avait de la compagnie, c'était assis sur le **front-porch**. (GRA:9/15)
(19) Ça c'est le **half-brother** à Mom ça. (TEM:4/61)

• **Les particules**. Cette catégorie regroupe, entre autres, des mots qui fonctionnent hors syntaxe ayant pour function de structurer le discours et d'organiser l'interaction. Donc *bien* en position préphrastique ne sera pas compté comme adverbe mais plutôt comme particule. Nous comptons les particules avec les mots grammaticaux.

(20) J'ai vu ton char à toi et j'ai vu un autre char, j'ai pensé **Lord** eyoù ce que Beverly est (ROD:10/28)
(21) **Boy** sa pauvre mère (BRE:10/13)
(22) Tu crois que c'est pour ça que t'as fait ça **like** pas parler français avec moi? (ROD:37/33)
(23) 1. Louise te parlait en français?
2. **Ah well** c'est pour ça que je je crois je l'ai appris un petit peu. (ROD:38/33)
(24) Elle voulait remplir un changement d'adresse, **all right,** ça se fait on avait la forme (GRA:10/16)
(25) On se parle en français, **but you see**, droite là, ma voisine droite là, ça c'était élevé français. (ROD:1/37)
(26) Mais au bord là, **oh yes** là où ce qu'on a été élevé. (TEM:21/62)
(27) **So** je crois qu'on respect each other privacy. (TEM:3/4)

• **Les verbes.**

(28) Tu crois que noncle Pierre a **enjoy** sa visite? (ROD:32/32)
(29) J'ai **retire** en 83 de chez McDermid (BRE:14/13)
(30) On a **ride** plein avec Fils (BRE:29/14)
(31) Ils ont **load** dessus les bateaux. (GRA:15/33)
(32) Ils ont **vandalize** l'école ouais. (BRE:146/24)
(33) J'ai **draw** mon security check. (BRE:14/13)
(34) Ça fait qu'on a **move** dedans un tas de différentes maisons. (TEM:25/63)
(35) Je crois ici là on peut mieux **discipline** notre enfant. (TEM:6/24)
(36) C'est là où elle s'est retrouvée quand Tom a **hang**, elle a fait appelé la loi, ça buvait à même les chars. (MOU:12/100)
(37) On grouillait avec le bateau à peu près trois cent pieds. On **feedait** de electric lines. (BRE:101/11)
(38) Ils sont **improuvés**, on fait une meilleure vie. (GRA:31/20)
(39) Ça se fait, ça **use** plus ce mot américain. (GRA:28/39)

**• La forme verbale *gone*.**

(40) Ils ont fait les cadiens **gone** de l'Acadie. (GRA:13/33)
(41) J'étions **gone** à la ville (GRA:16/29)
(42) La langue est après **gone**, l'influence du français était proche **gone**. (GRA:2/12)

**• Les verbes+prépositions**

(43) Mais quand ma mère vivait, elle pouvait **get by** tu sais.
(44) J'ai **go ahead** à faire pareil. (MOU:4/87)
(45) Sonya elle dit 'comment est-ce que tu dis welcome en français'? Je dis 'je sais pas mais je vas **find out** pour toi'. C'était pas de tracas mais j'allais **find out.** Ça fait que c'était bienvenue. (BRE:108/21)
(46) On était les seuls ici but là ça **build up** tout partout. (ROD:4/42)

**• Les conjunctions**

(47) Moi j'étais là un tas de fois **so** je pense que j'ai ramassé comme ça (ROD:39/33)
(48) Ils aimaient mieux ça là **because** le plus vieux, le plus elder monde au Pierre Part ça c'est: ça c'est juste français ça (ROD:41/33)
(49) Ils vont mettre leurs oignons, leurs tomates et tout **but** ils font pas un roux (BRE:92/19)

**• Les adjectifs**

(50) Le plus **elder** monde au Pierre Part ça c'est: ça c'est juste français ça (ROD:42/33)
(51) j'aime faire des gâteaux et ça et du **good old** bonne sauce à la volaille. (BRE:56/16)
(52) je l'avais vue en dehors mais son telephone était **busy**. (BRE:131/22)
(53) Ça te fait un effet que t'aimes pas à être **rude** à eux. (BRE:135/23)
(54) C'était pas **nice** tu sais. (BRE:92/11)
(55) Je pars pour la **whole** journée pour aller à Baton Rouge. (ROD:55/65)
(56) Tracy va pas parler tout à fait autant, ça c'est notre **second** garçon. (BRE:15/3)
(57) L'autre était manière **shy**, pas comme moi du tout. (ROD:45/56)
(58) Il a été **seasick** tu sais. (BRE:94/11)
(59) On find out s'ils faisont quelque chose de **wrong**. (TEM:46/65)

• **Les adverbes**

(60) Il drivait 24 heures **straight through** et là il restait... (BRE:7/13)
(61) **Now** moi si je vois les enfants des fois, je vais dire quelque chose en français. (BRE:24/4)

• **Les prépositions**

(62) On visitait notre cousin **across** la baie. (BRE:46/15)
(63) Je me tracasse pas trop **about** quand il va sortir et tout. (ROD:78/56)

• **Les verbes (français)+*back***

(64) Quand ils quittent, quand ils partent, ils reviennent toujours **back**, c'est leurs racines (ROD:1/5)
(65) Ils sont tous après ramener ça **back** asteur. Ils aiment que les enfants parlent en français (ROD:23/31)
(66) Quel est ton premier souvenir de la vie ici à Pierre Part? Est-ce que tu as:: tu peux penser **back** à ça? (ROD:45/34)
(67) Je viens **back** à la guerre, c'était pas **about** ça ici. (BRE:57/8)

| Catégorie de mots | N | % |
|---|---|---|
| Nom | 319 | 49.5 |
| Particule (*tag*, marque préphrastique, interphrastique) | 171 | 26.6 |
| Verbe | 68 | 10.5 |
| Conjonction | 38 | 5.9 |
| Adjectif | 19 | 3.0 |
| Adverbe | 14 | 2.2 |
| Préposition | 15 | 2.3 |
| Pronom | 0 | 0 |
| Déterminant | 0 | 0 |
| Total | 644 | 100 |

Tableau 2. La distribution grammaticale des emprunts

Le tableau 2 montre la répartition des emprunts à l'anglais selon la catégorie grammaticale. On voit que la catégorie "particule" représente plus de 25% des emprunts. Notre définition inclusive de cette catégorie influence forcément son usage important. À l'exception de cette catégorie discursive, la répartition des emprunts ressemble à celle de plusieurs autres études (ex: Poplack et al. 1988,

Poplack 1989, Flikeid 1989). On emprunte d'abord des noms, ensuite des formes verbales et d'autres mots lexicaux et, finalement, des mots grammaticaux. Nous n'avons trouvé aucun pronom ou déterminant emprunté à l'anglais.

**4. Assimilation dans la langue des mots des différentes catégories grammaticales.** Nous avons distingué deux types d'emprunts: les emprunts spontanés *"nonce loans"* et les emprunts établis. Ces derniers constituent des mots que nous avons pu retrouver dans le dictionnaire cadjin de Daigle (1984) et dans tous les lexiques (entre autres de nombreuses thèses) de français cadjin. Les emprunts repérés dans plus d'une entrevue ont également été classés comme des emprunts établis. Le tableau 3 montre que les particules sont généralement des emprunts établis. Bien que les catégories "préposition/conjonction" (regroupées à des fins d'analyse) s'empruntent moins souvent, elles se répandent rapidement une fois empruntées, tout comme les particules. Les mots lexicaux ont plus tendance à s'emprunter spontanément, quoique la majorité consiste en emprunts établis, surtout les noms.

| Catégorie | Récurrence | | | | Total |
|---|---|---|---|---|---|
| | Etabli | | Spontané | | |
| | N | % | N | % | |
| Particule | 169 | 99 | 2 | 1 | 171 |
| Prép./ Conj. | 46 | 87 | 7 | 13 | 53 |
| Noms | 240 | 75 | 79 | 25 | 319 |
| Adverbe/Adjectif | 20 | 60 | 13 | 40 | 33 |
| Verbe | 40 | 59 | 28 | 41 | 68 |
| Total | 515 | 80 | 129 | 20 | 644 |

Tableau 3. Récurrence versus catégorie grammaticale.

**5. Intégration phonologique des mots de différentes catégories grammaticales.** Le tableau 4 montre que les emprunts lexicaux sont plus intégrés phonologiquement que les mots grammaticaux. Bien que ce résultat ne surprenne guère - on s'attend à ce que les mots grammaticaux résistent à l'emprunt et à tout processus qui en découle, on se serait attendu (cf. Poplack et al. 1988) à ce que les emprunts récurrents, les mots grammaticaux dans le tableau 3, soient plus assimilés que les emprunts spontanés, qui se trouvent parmi les mots lexicaux selon le tableau 3. Nous examinerons cette question plus en détail dans la section 8.

| Catégorie | Intégration phonologique | | | | | | indét. | Total |
|---|---|---|---|---|---|---|---|---|
| | Totale | | Partielle | | Aucune | | | |
| | N | % | N | % | N | % | N | |
| Particule | 5 | 3 | 2 | 1 | 162 | 95 | 2 | 171 |
| Prép./Conj. | 1 | 2 | 2 | 4 | 49 | 92 | 1 | 53 |
| Adverbe/Adjectif | 2 | 6 | 2 | 6 | 29 | 88 | 0 | 33 |
| Verbe | 8 | 12 | 8 | 12 | 49 | 72 | 3 | 68 |
| Nom | 60 | 19 | 42 | 13 | 209 | 66 | 8 | 319 |
| Total | 76 | 12 | 56 | 9 | 498 | 77 | 14 | 644 |

Tableau 4. Intégration phonétique versus catégorie grammaticale.

**6. Intégration morphologique des noms.** En francais, le morphème du pluriel **-s** n'est pas prononcé dans un contexte de non-liaison. Par contre, il est prononcé en anglais. Pour être classé comme intégré, un emprunt à l'anglais devrait afficher la morphologie du français. Aucun des **-s** dans des contextes de non-liaison et aucun des autres marques de pluriel anglais ne devrait être prononcé si le mot est intégré au français. Cependant, le tableau 5 indique qu'une proportion de 28% des emprunts en contexte pluriel se produisent avec un [s] final. Mais on ne peut pas conclure qu'il s'agit du morphème de pluriel puisque les noms en contexte singulier affichent eux aussi un taux de 21% de [s] prononcé! Les noms d'origine anglaise en français cadjin s'empruntent accompagnés d'un [s] final dans 25% des cas, *indépendamment du nombre* du nom. La présence de [s] sur un emprunt à l'anglais ne représente donc pas une flexion du pluriel en langue anglaise.

| Morphologie | Ø | | -s | | |
|---|---|---|---|---|---|
| Nombre | N | % | N | % | Total |
| singulier | 177 | 79 | 48 | 21 | 225 |
| pluriel | 47 | 72 | 18 | 28 | 65 |
| Total (+ 29 indét.) | 224 | 77 | 66 | 23 | 290 |

Tableau 5. Intégration morphologique des noms.

**7. Intégration morphologique des verbes.** À l'exception des formes irrégulières comme *gone* ou verbe (anglais) + particule, les formes verbales prennent la morphologie temporelle du français. Dans le tableau 6, cinq occurrences sur neuf au présent étaient à la troisième personne singulière (exemple 39) et auraient pris le

morphème d'accord **-s** s'ils avaient été émis en anglais; aucune d'entre elles n'en possédait un. L'exception se trouve dans les formes non-temporelles d'origine anglaise, les infinitifs (exemple 35) et les participes passés (exemple 34) . On remarque une forte tendance à utiliser une forme qui correspond à l'infinitif anglais sans la particule *to*. La régularité de cette construction semble indiquer qu'il s'agit d'une stratégie spécifique au français cadjin. Il ne peut être question ici d'alternance de code puisqu'il n'y a aucune flexion anglaise de personne, de nombre ou de temps sur ces formes. Par exemple, on ne trouve ni le morphème de la troisième personne du singulier ni les formes irrégulières de l'anglais (*rode*: exemple 30, *drew*: exemple 33). L'interprétation la plus vraisemblable de l'utilisation de ces formes est qu'en empruntant une forme verbale à l'anglais dans un contexte où le français n'exige pas de marque de temps, on utilise la forme infinitive sans *to* et sans **-er**. Le fait que d'autres variétés du français n'affichent pas cette construction s'explique par les différentes conventions de classification des emprunts à l'intérieur des conjugaisons de verbes. Au lieu de se servir du morphème du participe passé **-é** (et **-er** avec l'infinitif), ce qui assignerait l'emprunt à la "première conjugaison", on ajoute Ø en français cadjin et, par conséquent, on établi une conjugaison spécifique aux emprunts. D'ailleurs cette règle pour le participe passé diffère peu de certaines autres conjugaisons (ex: *finir*). L'utilisation de la marque nulle sur le participe (et l'infinitif) d'un verbe emprunté à l'anglais ne représente donc pas une flexion verbale en langue anglaise.

| Temps/forme de verbe | +Morphologie française | | -Morphologie française | | Total |
|---|---|---|---|---|---|
| | N | % | N | % | |
| présent | 9 | 100 | 0 | 0 | 9 |
| infinitif | 3 | 25 | 9 | 75 | 12 |
| participe passé | 1 | 4 | 25 | 96 | 26 |
| imparfait | 4 | 100 | 0 | 0 | 4 |
| *gone* | 0 | 0 | 12 | 100 | 12 |
| verbe + particule | 0 | 0 | 5 | 100 | 5 |
| Total | 17 | 25 | 51 | 75 | 68 |

Tableau 6. Intégration morphologique des verbes.

**8. Intégration et récurrence.** Nous avons trouvé que les mots grammaticaux affichent un taux de récurrence élevé mais un taux

bas d'intégration phonologique. Selon nous (cf. Poplack et al. 1988), il serait plus logique que ces deux aspects de l'intégration d'emprunts aillent dans le même sens. Le tableau 7, qui compare l'intégration phonologique à la récurrence pour l'ensemble des emprunts, confirme cette hypothèse. Vingt-trois pour cent des emprunts récurrents sont totalement ou partiellement intégrés, alors que seulement 11% des emprunts spontanés sont intégrés.

| Intégration phonologique | Récurrence | | | | Total |
|---|---|---|---|---|---|
| | établi | | spontané | | |
| | N | % | N | % | |
| totale | 69 | 13 | 7 | 5 | 76 |
| partielle | 49 | 10 | 7 | 5 | 56 |
| aucune | 388 | 75 | 110 | 85 | 498 |
| indéterminée | 9 | 2 | 5 | 4 | 14 |
| Total | 515 | 100 | 129 | 99 | 644 |

Tableau 7. Intégration phonologique des emprunts versus assimilation dans la communauté.

Comment expliquer la contradiction entre la tendance attendue et observée dans le tableau 7 et les résultats des tableaux 3 et 4? Le tableau 8 montre la répartition des données du tableau 7 selon la catégorie grammaticale de l'emprunt. Pour chaque catégorie de mots lexicaux, le pourcentage d'intégration totale est plus élevé pour les emprunts établis que pour les emprunts spontanés. De plus, pour chaque catégorie de mots lexicaux, le pourcentage d'intégration partielle est plus élevé pour les emprunts établis que pour les emprunts spontanés. Ces données sont conformes à nos attentes. Or, il n'y a pas assez de cas (N=13) pour vérifier si les mots grammaticaux empruntés de façon spontanée sont intégrés. Donc la corrélation entre l'intégration phonologique et la récurrence se confirme à l'intérieur des catégories grammaticales qui contiennent suffisamment de données. Dans le cas de mots grammaticaux, cette corrélation n'est pas pertinente parce que les emprunts spontanés sont rares. Il y a peu de variation; presque tous les emprunts de ce type sont déjà mentionnés dans la littérature sur le français cadjin, et il n'y a pas d'évidence d'intégration à la phonologie française.

**Conclusion**. Nous ne prétendons pas que notre corpus soit représentatif, statistiquement parlant, de la langue cadjine en Louisiane. Cependant les résultats sont assez clairs pour nous permettre d'établir des hypothèses de départ pour une future étude à partir d'un corpus dont la construction est présentement en cours à Louisiana State University.

Plusieurs observations importantes se dégagent de notre étude. Tout d'abord le taux d'emprunt, 1,6 occurrences par cent mots de discours, en français cadjin n'est pas plus élevé que les taux cités pour d'autres variétés nord-américaines du français.

| Intégration phonologique | Récurrence | | | | Total |
|---|---|---|---|---|---|
| | établi | | spontané | | |
| | N | % | N | % | |
| Particule/Préposition/ Conjonction | | | | | |
| totale | 6 | 3 | 0 | 0 | 6 |
| partielle | 2 | 1 | 2 | 22 | 4 |
| aucune | 204 | 95 | 7 | 78 | 211 |
| indéterminée | 3 | 1 | 0 | 0 | 3 |
| total | 215 | 100 | 9 | 100 | 224 |
| Nom | | | | | |
| totale | 55 | 23 | 5 | 6 | 60 |
| partielle | 38 | 16 | 4 | 5 | 42 |
| aucune | 143 | 60 | 66 | 84 | 209 |
| indéterminée | 4 | 2 | 4 | 5 | 8 |
| total | 240 | 101 | 79 | 100 | 319 |
| Verbe | | | | | |
| totale | 6 | 15 | 2 | 7 | 8 |
| partielle | 7 | 18 | 1 | 3 | 8 |
| aucune | 25 | 63 | 24 | 86 | 49 |
| indéterminée | 2 | 5 | 1 | 3 | 3 |
| total | 40 | 101 | 28 | 99 | 68 |
| Adjectif/Adverbe | | | | | |
| totale | 2 | 10 | 0 | 0 | 2 |
| partielle | 2 | 10 | 0 | 0 | 2 |
| aucune | 16 | 80 | 13 | 100 | 29 |
| indéterminée | 0 | 0 | 0 | 0 | 0 |
| total | 20 | 100 | 13 | 100 | 644 |

Table 8. Intégration versus assimilation, selon la catégorie grammaticale.

On dit que les Cajuns font plus de mélange de langue quand ils parlent entre eux. Quoiqu'il ne s'agisse pas d'un grand échantillon d'entrevues, ce n'est pas le cas pour les deux entrevues de ce genre dans notre corpus, en comparaison avec les trois entrevues dont l'intervieweur est un non-membre qui parle une variété de français apprise à l'école ("le français standard").

Les mots empruntés à l'anglais ne sont pas tirés du lexique au hasard. Si on compte les emprunts différents, il y a une nette prépondérance de noms et d'autres mots lexicaux et relativement peu de mots grammaticaux, comme en témoignent de nombreuses autres études (ex: Poplack et al. 1988), quoique certains mots grammaticaux empruntés reviennent souvent. Nous avons repéré certaines constructions calquées sur l'anglais, telles que les constructions verbe (français) + préposition (anglais), dont la plus répandue est *back* (exemples 64-67), ainsi que l'utilisation fréquente de la forme invariable *gone* comme participe ou adjectif dans plusieurs contextes. Ces constructions sont très intéressantes parce qu'elles sont également répandues dans le français acadien de la Nouvelle Ecosse (Flikeid 1989) et dans le français du Nouveau-Brunswick (Péronnet 1989), mais inexistantes ou peu utilisées au Québec.

Nous avons trouvé que pour les mots lexicaux, le pourcentage des emprunts totalement ou partiellement intégrés phonologiquement est plus élevé pour les emprunts établis que pour les emprunts spontanés. Les mots grammaticaux semblent se comporter différemment. Ceux-ci sont généralement résistants à l'emprunt. Cependant, selon nos définitions opératoires, il y a quelques centaines d'occurrences d'emprunts à l'anglais dans cette classe de mots. De plus, ces emprunts sont établis mais peu intégrés. Il y a plusieurs explications possibles pour ces résultats. Il s'agit souvent des mots préphrastiques, comme des particules d'interaction ou des conjonctions, tandis que les emprunts de mots lexicaux ont lieu presque toujours à l'intérieur des phrases, et même à l'intérieur des groupes nominaux et des groupes verbaux. En début de phrase ou de tour de parole, est-ce qu'on est plus susceptible de subir l'interférence linguistique? Ainsi ces occurrences seraient le résultat de l'interférence ou de l'alternance de langue, plutôt que l'emprunt. Ceci expliquerait pourquoi ces occurrences ne manifestent pas l'intégration phonologique. Et le fait que cette classe de mots est fermée expliquerait pourquoi les mêmes mots se répètent souvent.

D'autres explications se baseraient sur le bilinguisme de nos locuteurs, sur le manque de sites possibles pour l'intégration phonologique dans les mots courts comme *but* et *so*, ou sur le besoin d'une grille d'analyse phonétique plus fine pour détecter l'intégration dans ces mêmes mots. Ces hypothèses, et d'autres, seront mises à l'épreuve lors de notre recherche projetée sur un plus grand corpus.

Les résultats les plus intéressants portent sur des aspects morphologiques des emprunts nominaux et verbaux. Dans les deux cas, on pourrait penser, à première vue, qu'il n'y a pas d'intégration morphologique en français et que c'est la morphologie de l'anglais qui s'applique. Néanmoins, dans le cas de la morphologie du pluriel, nous avons montré que la présence du **-s** prononcé est plutôt reliée à une tendance réelle (autour de 25%) à choisir la forme plurielle du mot anglais, indépendamment du contexte (singulier ou pluriel) dans le discours français où l'emprunt se trouve. Dans le cas des formes atemporelles des verbes, infinitifs et participes passés, nous avons montré qu'il n'y avait aucune trace de la morphologie anglaise. La racine de l'infinitif du verbe anglais est utilisée de façon assez stricte après l'auxiliaire français dans les temps composés sans aucune indication morphologique de temps, de nombre ou de personne en anglais, et sans le *to* dans le cas de l'infinitif.

**Bibliographie**

Daigle, Jules. 1984. *A Dictionary of the Cajun Language in 1984* Ville Platte/LA: Swallow Publications.

Dubois, Sylvie. 1995. "Field methods in Cajun communities". Dans A. Valdman (réd.) *French and Creole in Louisiana*. Indiana: Penblum.

Dubois, Sylvie, William Gautreau, Howard Margot & Tracy Veler. 1995. "Laissez le français rouler". Dans *Proceedings of the 3rd Annual Symposium on Language and Society*. Austin: University of Texas-Austin, Texas Forum, pp.161- 175.

Dubois, Sylvie, William Gautreau, Howard Margot, Megan Melançon & Tracy Veler. 1995. "The quality of French spoken in Louisiana". *SECOL Review* 19,2:16-39.

Dubois, Sylvie & Megan Melançon. 1996. "Cajun is dead: Long live Cajun." *Journal of Sociolinguistics* 1,1.

Flikeid, Karin. 1989. "Moitié anglais, moitié français. Emprunts et alternance de langues dans les communautés acadiennes de la Nouvelle-Écosse." *Revue québécoise de linguistique théorique et appliquée* 8,2: 177-228.

Meechan, Marjory & Shana Poplack. 1995. "Orphan categories in bilingual discourse: A comparative study of adjectivization strategies in Wolof/French and Fongbe/French". *Language Variation and Change* 7: 169-194.

Pellerin, Eveline. 1937. *La langue française en Louisiane*. Thèse de doctorat, Université McGill.

Péronnet, Louise. 1989. "Analyse des emprunts dans un corpus acadien." *Revue québécoise de linguistique théorique et appliquée* 8,2:229-252.

Poplack, Shana. 1989. "Statut de langue et accommodation langagière le long d'une frontière linguistique." Dans R. Mougeon et E. Beniak (réd.) *Le français parlé hors Québec*. Québec: Les Presses de l'Université Laval, pp. 127-151.

Poplack, Shana, David Sankoff & Chris Miller. 1988. "The social correlates and linguistic processes of lexical borrowing and assimilation." *Linguistics* 26: 47-104.

Poplack, Shana & Marjory Meechan. 1995. "Patterns of language mixture: nominal structure in Wolof-French and Fongbe-French bilingual discourse". Dans P. Muysken & L. Milroy (réd.) *One speaker, two languages*. Cambridge: Cambridge University Press. pp.199-232.

Sankoff, David, Shana Poplack & Swathi Vanniarajan. 1990. "The case of the nonce loan in Tamil." *Language Variation and Change* 2: 71-101.

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 177-188.

# Analyse du contexte d'utilisation de quelques emprunts en vue de leur désambiguïsation automatique

*Inès Escayola, Isabelle Lavigne, Michel Théoret*
Université de Sherbrooke

## INTRODUCTION

Dans le cadre des travaux du CATIFQ (Centre d'analyse et de traitement informatique du français québécois), l'un des sous-projets[1] s'intéresse plus particulièrement au phénomène des anglicismes, le mot étant entendu ici en son sens restrictif d'emprunt dont l'emploi est critiqué parce que considéré comme abusif ou inutile. L'objectif de ces travaux est d'analyser d'une part la fréquence de ces anglicismes, et plus particulièrement leur fréquence relative selon les types de discours, d'autre part leur contexte d'emploi, le tout devant servir de base à la description prédictionnairique de cette tranche particulière du vocabulaire québécois. Pour y arriver, comme nous l'annoncions lors d'un précédent colloque de l'ACFAS (Cajolet-Laganière et Théoret 1994), nous avons regroupé tous les anglicismes donnés par les sources québécoises les plus fournies à cet égard (les chroniques et répertoires de l'OLF, les deux premières éditions des *Anglicismes* de Colpron, le *Multidictionnaire* de Marie-Éva de Villers, le *Dictionnaire des canadianismes* de Dulong), auxquels nous avons ajouté ceux qui étaient déjà ainsi étiquetés dans le corpus oral de Sherbrooke (Beauchemin, Martel et Théoret, 1992); nous avons rangé ces mots dans un dictionnaire électronique en notant toutes les formes sous lesquelles ils peuvent se rencontrer et en les codifiant selon six catégories, allant des mots simples non ambigus jusqu'aux groupes complexes calqués et tout à fait ambigus. Cette partie du travail est terminée; dans l'étape suivante, actuellement en cours, nous mettons au point un logiciel de repérage permettant de balayer, à l'aide de ce dictionnaire électronique, l'ensemble des textes de la BDTS (Banque

[1] L'équipe de recherche comprend Hélène Cajolet-Laganière, Pierre Martel et Michel Théoret ainsi que plusieurs assistantes, notamment Inès Escayola, Isabelle Lavigne et Sylvie Thiboutot.

de données textuelles de Sherbrooke)[2] contenant à ce jour un peu plus de cinq millions de mots.

Nous présentons ici l'état des travaux en commentant, dans une première partie, quelques chiffres généraux et les résultats préliminaires des premiers balayages effectués sur les formes non ambiguës; en seconde partie, nous proposerons une analyse de quelques mots ambigus, pour voir comment arriver à la désambiguïsation la plus poussée possible.

## 1. PREMIÈRE PARTIE : CHIFFRES GÉNÉRAUX ET RÉSULTATS DES PREMIERS BALAYAGES

### 1.1. Quelques chiffres généraux

Avant de passer aux résultats proprement dits, précisons que le dépouillement des sources a été fait systématiquement, sans donc porter de jugement sur le bien-fondé de l'étiquette « anglicisme »; cela ne pourra être fait qu'après avoir recueilli toutes les données d'analyse. De plus, il reste quelques corrections à apporter, qui nous ont été indiquées par les premiers balayages; dans l'interprétation des chiffres que nous livrons ici, ces variables doivent être prises en compte.

#### 1.1.1. Répartition selon le type et le degré d'ambiguïté des vocables

Une fois le dépouillement terminé, le dictionnaire électronique contient **4227 vocables** ou 10 682 formes[3], qui se répartissent en catégories :

---

[2] Pour plus de renseignements sur le contenu de la BDTS, on peut s'adresser, par internet, à : http://www.usherb.ca/catifq/banque.html#bdts

[3] Les formes comprennent évidemment toutes les variantes morphologiques, pluriels, féminins, formes conjugées ainsi que les variantes graphiques (*chum / tchome, fun / fonne, phonne, etc.*).

**A. les non ambigus**

| | **Vocables** | **Formes** |
|---|---|---|
| 1.mots simples de forme anglaise (*anyway, chum*, etc.) | 1821 | 4848 |
| 2.groupes complexes, dont un terme non ambigu (*à la botch*) | 701 | 1576 |
| 3.groupes complexes, de type calque non ambigu (*hors d'ordre*) | 518 | 1310 |
| **Total des non ambigus** | **3040** | **7734** |

**B. les ambigus**

| | **Vocables** | **Formes** |
|---|---|---|
| 4.mots simples, homographes de mots français (*tire, date* (rendez-vous)) | 192 | 427 |
| 5.mots simples de type calque (*pamphlet, régulier*) | 758 | 2091 |
| 6.groupes complexes ambigus (*manquer le bateau*) | 237 | 430 |
| **Total des ambigus** | **1187** | **2948** |

Les **non ambigus** représentent donc **72 %** et des vocables et des formes; les **ambigus, 28 %.**

### 1.1.2. Répartition selon les classes grammaticales

Dans notre dictionnaire électronique ainsi constitué, les classes grammaticales se répartissent comme suit :

| | |
|---|---|
| les **substantifs** forment | 73 % de l'ensemble |
| les **adjectifs** | 14 % |
| les **verbes** | 11 % |
| le **reste** (adverbes, etc.) | 2 % |

Il n'est évidemment pas étonnant que les substantifs soient en surnombre; la quantité de verbes est cependant un peu inférieure à ce qu'on pouvait attendre, compte tenu des conclusions auxquelles nous sommes déjà arrivés au cours de recherches parallèles (Théoret, 1993). Cet écart peut s'expliquer d'une part par la différence entre discours oral et langue générale, ce que nous pourrons vérifier bientôt, d'autre part et surtout par le caractère théorique du dictionnaire électronique; c'est pourquoi nous dirons maintenant un mot des résultats des premiers balayages que nous avons effectués dans la BDTS.

### 1.2. Le balayage des mots ou groupes non ambigus

Au strict plan du balayage des textes, les mots ou groupes de mots non ambigus ne posent pas beaucoup de problèmes; les mots simples se cherchent directement dans la liste de base, puisque celle-ci contient déjà toutes les formes possibles. Mais nous avons quand même simplifié le travail en introduisant un conjugueur automatique pour les formes verbales ou en faisant faire une partie du balayage par de simples radicaux quand la chose était possible; nous avons aussi raffiné la recherche en introduisant quelques règles de distinction automatique entre des formes homographes (formes verbales et substantives, par exemple, souvent distinguées par le mot qui précède). Sans expliquer dans le détail le fonctionnement du logiciel de repérage, disons tout de même que, pour les groupes complexes (et ceci vaut, quel que soit le degré d'ambiguïté du groupe en question), il fallait d'abord déterminer le mot de base. Nous devions de plus prévoir la longueur du contexte dans lequel faire la recherche dans

tous les cas d'expressions non complètement figées où des mots étrangers peuvent interférer (par exemple, dans *aller en appel* ou *parler à travers son chapeau*, divers adverbes peuvent se glisser entre le verbe et le complément, *il n'ira certainement pas en appel; il parle, de façon constante, à travers son chapeau...*). Ces façons de procéder mènent même à la désambiguïsation de certains éléments apparemment tout à fait ambigus (des catégories 4, 5 et 6); et l'étude des contextes permettra, pour ces derniers, de pousser encore plus loin l'automatisation du repérage.
Nous commenterons maintenant quelques résultats du balayage des mots simples non ambigus.

### 1.2.1. Le balayage des mots simples non ambigus

Ces mots constituent, rappelons-le, le groupe le plus nombreux de notre dictionnaire électronique, avec 1821 vocables et 4848 formes, ce qui constitue respectivement 43 % et 45 % de l'ensemble.

### 1.2.2. Les vocables absents

Il est tout d'abord intéressant de noter que près de la moitié (882 / 1821) des vocables du dictionnaire électronique n'ont aucune occurrence dans notre corpus de plus de 5 millions de mots. La proportion paraît énorme, et l'explication du phénomène peut relever de plusieurs ordres :

- premièrement, à 5 millions de mots, le corpus de Sherbrooke est encore relativement petit; l'équipe du CATIFQ travaille à le développer tout en améliorant sa représentativité, aspect qui nous paraît d'ailleurs plus important que la quantité d'occurrences en soi. Nous avons en outre plusieurs autres tranches de vocabulaire non encore intégrées à la BDTS, de même que des corpus comme ceux du CD-ROM Actualités servant, à cette étape-ci, de corpus-témoins que nous pouvons faire balayer par le logiciel de repérage, ce qui permet de confirmer ou d'infirmer nos résultats;
- deuxièmement, les sources dont nous avons tiré notre dictionnaire forment une sorte de liste théorique, contenant des mots dont l'usage est disparu ou en train de disparaître (*théâtre* au sens de

cinéma, *intercom*, des mots relatifs aux techniques anciennes des bûcherons ou des cultivateurs, souvent cités par Dulong), contenant aussi des termes de langue de spécialité (droit, assurances, informatique) que nous ne trouverons que dans des corpus particuliers.

Il reste que l'on peut être surpris de l'absence de certains, *schedule*, *secure*, *brainstorming* ou *cancellation*. À l'inverse, ces sources étant en quelque sorte figées, au moins à la date de leur dernière édition, on n'y trouve pas un certain nombre de termes plus récents et très à la mode comme *top model* ou *compact disc*. Nos travaux ultérieurs devront évidemment tenir compte de tous ces phénomènes, mais notons que, déjà, nous ajoutons les « nouveaux » anglicismes au dictionnaire au fur et à mesure que nous les relevons.

### 1.2.3. Les mots simples présents

Les mots simples non ambigus représentent, dans le corpus de 5 millions de mots, 12 938 occurrences rangées sous 939 vocables. Il n'y a guère d'enseignements à tirer de ces chiffres globaux, mais nous pouvons encore regarder les résultats des trois tranches les plus importantes; les corpus de langue orale, de langue journalistique et de langue littéraire (cf. note 2), même s'ils ne forment pas des ensembles parfaitement égaux, sont quand même comparables, se situant entre 1,4 et 1,6 millions d'occurrences. Les anglicismes de la première catégorie y sont distribués de la manière suivante :

| | **Langue orale** | **Langue journalistique** | **Langue littéraire** |
|---|---|---|---|
| occurrences | 5527 | 4424 | 1809 |

Que les anglicismes soient plus nombreux en langue orale n'a rien de bien étonnant. Que leur nombre soit presque aussi élevé dans la langue journalistique, et beaucoup plus qu'en langue littéraire, l'est toutefois davantage. Mais il ne faut pas en tirer de conclusions trop hâtives, et il sera sûrement intéressant de voir si ce sont les mêmes emprunts qui se trouvent dans les trois tranches, ainsi que leur importance dans chacune. Pour en avoir une idée, comparons les occurrences des trois

anglicismes qui, dans le *Dictionnaire de fréquence* (Beauchemin et coll. 1992) de langue orale avaient la plus haute fréquence, soit *fun, job* et *chum* :

| | **Langue orale** | **Langue journalistique** | **Langue littéraire** |
|---|---|---|---|
| ***fun*** | 333 | 19 | 47 |
| ***job*** | 149 | 41 | 39 |
| ***chum*** | 151 | 8 | 21 |

Ce tableau fait ressortir clairement que la distribution des anglicismes, de ce type en tout cas, n'est pas du tout la même entre trois tranches relativement égales. Cette conclusion est fort intéressante dans la mesure où elle montre qu'une étude des indices de fréquence, de dispersion et d'usage permettrait, entre autres choses, d'établir des bases solides pour déterminer les marques à leur donner dans un éventuel dictionnaire.

Mais nous n'en sommes pas encore là. Et, pour que le travail soit complet, il faut encore que le logiciel de repérage soit tout à fait au point, ce qui pose des problèmes de désambiguïsation sur lesquels nous allons maintenant nous pencher au moyen de quelques exemples.

## 2. DEUXIÈME PARTIE : CAS AMBIGUS

En ce qui concerne les catégories d'anglicismes dits « ambigus », le travail consiste à balayer la BDTS et à relever tous les contextes dans lesquels ces mots sont employés afin d'établir des processus de désambiguïsation automatique dans le plus grand nombre de cas possible.

Les anglicismes que nous considérons comme ambigus, on l'a dit, sont de trois sortes : les mots simples, homographes de mots français; les mots simples de type calque; les groupes de mots ambigus.

Aux fins de cet article, nous nous sommes concentrés sur la catégorie regroupant des mots simples français de type calque. Ils comprennent, on l'a vu, 2091 formes ou 758 vocables. En général, ce sont les mots qui causent le plus de problèmes aux rédacteurs et aux réviseurs. Dans

les dictionnaires généraux - français et québécois - ou dans les dictionnaires de difficultés de la langue, ils sont souvent traités de manière incomplète ou inappropriée au contexte québécois.

**2.1. Méthodologie**

Voici brièvement comment nous avons procédé pour cette étude. Nous avons d'abord choisi quatre anglicismes représentatifs de cette catégorie, soit *compléter, opérer, régulier* et *significatif*. Puis, nous avons retenu sept dictionnaires : deux dictionnaires généraux français (*Nouveau Petit Robert* et *Petit Larousse illustré*), trois dictionnaires généraux québécois (*Dictionnaire nord-américain de la langue française* (Bélisle), *Dictionnaire du français Plus* et *Dictionnaire québécois d'aujourd'hui*) et deux dictionnaires québécois de difficultés de la langue (*Dictionnaire des difficultés de la langue française au Canada* (Dagenais) et *Multidictionnaire des difficultés de la langue française*).

Pour chacun des quatre anglicismes, nous avons procédé à un recensement exhaustif de tous les sens et sous-sens donnés par les sept dictionnaires, puis à un relevé des sens critiqués.

Nous avons constitué une grille regroupant l'ensemble de ces sens, sous-sens et sens critiqués. Puis, nous avons balayé la BDTS pour recueillir tous les contextes où apparaissent les quatre mots. À l'aide de la grille, nous avons procédé à l'analyse de ces contextes. Nous avons tenté de voir si les sens relevés dans la BDTS pour les quatre anglicismes correspondaient aux sens et sous-sens listés dans les dictionnaires. Nous voulions vérifier également si les sens critiqués répertoriés par certains dictionnaires se retrouvaient effectivement dans les contextes de la BDTS. Enfin, si l'on retrouvait, dans ces contextes, des sens non répertoriés par les dictionnaires.

Cette analyse nous a en outre permis de relever certaines particularités pour chacun des quatre anglicismes traités. Par exemple, tous les dictionnaires généraux (français ou québécois) ne retiennent pas les mêmes sens et sous-sens; les dictionnaires généraux français ne mentionnent pratiquement jamais les sens critiqués des mots à l'étude; les dictionnaires québécois généraux ou de difficultés ne les relèvent

pas toujours; certains sens sont abondamment utilisés dans la BDTS, certains autres, jamais.

L'analyse a également permis de dresser, pour chacun des quatre anglicismes, une liste de cooccurrents; ceux-ci, parce qu'ils apparaissent régulièrement dans l'entourage du mot à l'étude, deviennent un outil précieux pour la désambiguïsation. Peu importe, par ailleurs, qu'ils se retrouvent ainsi avec un sens correct ou un sens critiqué. Selon le cas, le logiciel peut relever l'anglicisme lorsque le mot est accompagné de tel cooccurrent ou l'emploi correct s'il est utilisé avec tel autre cooccurrent; la désambiguïsation s'en trouve ainsi de toute manière facilitée.

L'étude des contextes d'utilisation de ces quatre anglicismes nous a permis de cerner un certain nombre de cas et d'exemples où la désambiguïsation automatique est facile et réalisable, et d'autres, où elle est plus difficile, voire irréalisable. Voici, de façon globale, ce que nous avons recueilli pour chacun de ces anglicismes.

### 2.2. L'adjectif *significatif*

L'adjectif *significatif* semble relativement facile à désambiguïser. Le balayage du mot dans la BDTS nous donne 36 cooccurrents différents de l'adjectif *significatif*. Parmi ces cooccurrents, certains sont toujours employés dans les sens corrects, d'autres, toujours dans les sens incorrects. Ils présentent donc peu d'équivoques. Ainsi, on pourrait programmer le logiciel pour qu'il relève l'adjectif *significatif* lorsque celui-ci est accompagné de « coûts », de « frais », de « sommes », de « progrès », etc., cas pour lesquels l'analyse des contextes a démontré qu'il s'agissait le plus souvent d'emplois critiqués. Au contraire, le logiciel ne relèverait pas les *significatifs* précédés de « écarts », « résultats », « vote », « chiffres », « sondage », « symbole », « geste », « regard », etc., car, en général, ces expressions sont employées correctement dans les contextes de la BDTS.

### 2.3. Le verbe *opérer*

De même, le verbe *opérer* semble relativement facile à désambiguïser. Le balayage nous donne 16 cooccurrents différents. Lorsque ces

cooccurrents concernent la médecine, la chirurgie, lorsqu'ils expriment une partie du corps (*opérer le foie*) ou une maladie (*opérer un cancer, une tumeur*), le verbe *opérer* est bien utilisé : le logiciel devrait être en mesure d'*opérer* efficacement. D'autant plus que la forme pronominale est souvent utilisée avec ces cooccurrents (*se faire opérer, se faire opérer pour le foie*). À l'inverse, lorsque les cooccurrents désignent des outils ou des appareils, on programmerait le logiciel pour qu'il ne relève que des emplois critiqués. Quant aux contextes de « opérer un commerce » ou « opérer une entreprise », ils sont faciles à désambiguïser, pour les mêmes raisons.

### 2.4. L'adjectif *régulier*

En ce qui concerne l'adjectif *régulier*, l'analyse des contextes d'utilisation (qui présentent 39 cooccurrents différents), de même que le dépouillement des exemples donnés par les dictionnaires, révèle qu'il est souvent précédé des mots « café », « essence », « séance », « prix », etc. Le dépouillement des dictionnaires permet de déduire que ces emplois de l'adjectif *régulier* sont critiqués. On pourrait donc faire en sorte que le logiciel relève comme emplois critiqués tous les *réguliers* accompagnés de ces cooccurrents. À l'inverse, le logiciel aurait comme directive de laisser passer les *réguliers* accompagnés de « mouvement », de « traits », de « écriture », de « forme », etc., expressions où, de prime-abord, l'adjectif *régulier* est bien employé. Cet exercice d'apparence simple se complique cependant pour des groupes du genre « employé régulier » ou « horaire régulier », où le contexte est la seule référence pour décider du statut de l'expression. En effet, il suffit d'une légère nuance pour que l'adjectif passe de l'emploi correct à l'emploi critiqué. « Employé régulier » au sens d'« employé permanent » est qualifié d'incorrect, alors qu'au sens d' «employé assidu, ponctuel, consciencieux », il est tout à fait correct. De la même façon, « horaire régulier » est critiqué au sens de « horaire habituel » (par opposition à horaire d'été par exemple), alors qu'il est correct au sens de « horaire constant, qui ne varie pas » (toujours de 9 h à 5 h) par exemple. Ce genre d'équivoque revient dans un certain nombre de cas où le contexte devient vraiment

indispensable. C'est là que la désambiguïsation automatique se complique, puisque, il va sans dire, l'ordinateur est incapable d'analyser la valeur sémantique des contextes avec toutes ces nuances. Il faudra donc, dans ces cas, indiquer à l'utilisateur toutes les nuances d'emploi, avec des exemples concrets.

### 2.5. Le verbe *compléter*

*Compléter* présente le même type de difficulté que *régulier* dans le sens où les cooccurrents ne suffisent pas à la classification des utilisations : le contexte doit absolument être étudié pour trancher. Le logiciel peut cependant très bien s'occuper de cas comme « compléter un questionnaire » ou « compléter un formulaire », où, dans la presque totalité des contextes, il s'agit d'emplois critiqués.

### CONCLUSION

Voilà, dans les grandes lignes, ce qui ressort des premières analyses de ces quatre anglicismes. Nous allons procéder de la même façon pour tous les mots ou groupes de mots ambigus dans le but, comme nous l'avons dit, d'en arriver à la désambiguïsation la plus automatique possible. Quant à l'ensemble des travaux, il nous permettra, d'une part, de réaliser quelques études - quantitatives notamment - sur cette partie du vocabulaire québécois et, d'autre part, de contribuer utilement d'abord à sa description prédictionnairique, puis à l'établissement de marques d'usage dans un éventuel recueil d'anglicismes ou dans un dictionnaire de type général.

Par ailleurs, le logiciel de repérage que nous mettons au point pourra être mis à la disposition de rédacteurs et de réviseurs qui y trouveront un outil efficace pour balayer divers types de textes, repérer des emplois critiqués et trouver un certain nombre de solutions de rechange. Nous sommes conscients qu'il restera des cas où la désambiguïsation automatique ne sera pas possible; mais, grâce à ce travail d'analyse des contextes, nous serons en mesure d'aider l'utilisateur à effectuer lui-même, selon le type de discours sur lequel il travaille, le choix le plus approprié.

**BIBLIOGRAPHIE**

BEAUCHEMIN, Normand, Pierre MARTEL et Michel THÉORET. 1992. *Dictionnaire de fréquence des mots du français parlé au Québec. Fréquence, dispersion, usage, écart réduit*. New York, Peter Lang.

BÉLISLE, Louis-Alexandre. 1971. *Dictionnaire général de la langue française au Canada*. 2e éd. Québec, Bélisle Éditeur Inc.

BOULANGER, Jean-Claude, et coll. 1993. *Dictionnaire québécois d'aujourd'hui*. Montréal, Dicorobert.

CAJOLET-LAGANIÈRE, Hélène et Michel THÉORET. 1994. « Un dictionnaire électronique d'anglicismes ». Communication au Congrès de l'ACFAS. Polycopié.

COLPRON, Gilles. 1970 et 1982. *Les anglicismes au Québec*. Montréal, Beauchemin.

DAGENAIS, Gérard. 1984. *Dictionnaire des difficultés de la langue française au Canada*. Québec, Éditions françaises.

DE VILLERS, Marie-Éva. 1992. *Multidictionnaire des difficultés de la langue française*. Montréal, Québec-Amérique.

DULONG, Gaston. 1989. *Dictionnaire des canadianismes*. Montréal, Larousse.

*Petit Larousse illustré*. 1994. Paris, Larousse.

POIRIER, Claude. 1988. *Dictionnaire du français plus*. Montréal, CÉC.

REY-DEBOVE, Josette et Alain REY. 1993. *Le Nouveau Petit Robert*. Paris, Le Robert.

THÉORET, Michel. 1993. « L'emprunt en France et au Québec : à causes différentes, effets différents », dans Français de France-français du Canada. Coll. « Canadiana Romanica » n° 7, Tübingen, Niemeyer, p. 217-230.

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 189-200.

# Lexique bilingue et créativité: modèle des verbes en langue mixte dans une situation de contact entre l'arménien et le français

*Sylvia Kasparian*
Université de Moncton

*Résumé: Dans le présent article, nous nous proposons de décrire un modèle de verbes mixtes arméniens-français qui se retrouve dans le parler bilingue de Français d'origine arménienne vivant à Paris. Nous essayerons d'établir la nature ainsi que les moyens linguistiques qui ont servi à la création de ces nouveaux verbes bilingues.*

Une langue est en perpétuelle évolution, résultat d'un mouvement interne propre à la dynamique de toute langue ainsi que des facteurs extralinguistiques. Cette évolution et cet enrichissement de la langue se manifestent par la création de termes et de structures qui répondent à une nécessité, celle de traduire des situations ou des concepts nouveaux en recourant à la langue. Une situation de contact de langues est d'autant plus propice à ces créations qu'elle permet la rencontre de deux structures, deux idéologies, deux visions du monde véhiculées par deux cultures à travers deux langues différentes.

La créativité, le renouveau dans une langue reflètent ainsi l'environnement dans lequel elle évolue. En ce qui concerne l'arménien, la dispersion des Arméniens à travers le monde depuis le génocide de 1915 et leurs rencontres avec les langues des pays d'accueil ont beaucoup favorisé tous les phénomènes de contact de langues. L'apparition de ces nouveautés s'effectue de diverses manières, celles-ci étant en rapport direct avec la nécessité du nouveau terme ou de la nouvelle structure. Nous trouvons, le plus souvent, des cas d'emprunts lexicaux et de calques syntaxiques d'une langue à une autre qui, sous la pression de la fréquence et de la généralisation de l'utilisation, s'intègrent petit à petit à la langue arménienne.

Le nombre des emprunts est considérable dans la langue arménienne contemporaine. L'étude des emprunts et calques dans cette langue s'est déjà faite dans quelques situations de contact de langues, notamment en Arménie ex-soviétique.

Le présent article considère un phénomène de contact de langues qui, à notre connaissance, n'a presque pas été traité dans l'étude de la langue arménienne contemporaine et qui ne correspond ni à un calque ni à un emprunt, mais à un mélange de codes qui aboutit à la *création* d'une nouvelle structure en langue mixte. Contrairement à l'emprunt qui consiste en l'intégration (avec ou sans modification) d'éléments d'une langue au système linguistique d'une autre langue, c'est, là, le cas d'un mélange de deux systèmes, qui engendre un troisième système; l'unité empruntée n'est pas intégrée au système de la deuxième langue mais son traitement morpho-syntaxique dans la deuxième langue engendre un nouveau système. Le terme *création* est utilisé dans ce texte comme équivalent des mots *restructuration* ou *métissage* utilisés par d'autres linguistes (cf. Perrot (1995) Romaine (1995) et autres).

Le phénomène observé fait partie de la langue parlée; il constitue un cas de création d'une nouvelle structure verbale bilingue, arménien-français, qui est très fréquemment utilisée dans la langue courante d'Arméniens bilingues vivant à Paris[1]. Cette structure se révèle autant dans les conversations de locuteurs parlant la variété de l'arménien occidental (arméniens du Liban, de la France, des États-unis, etc.) que ceux parlant l'arménien oriental (arméniens d'Iran et d'Arménie).

À travers l'exposé de ce cas, nous nous efforcerons d'établir la *nature* de cette création et de déterminer les *moyens linguistiques* qui ont servi à l'élaboration de ce néologisme.

---

[1] Extrait d'un travail de recherche sur le parler bilingue-multilingue d'Arméniens vivant à Paris. Cf. Kasparian (1992).

Ces nouveaux verbes bilingues sont construits sur le modèle suivant:

| Lexème<br>Participe passé<br>(français) | + | Suffixe<br>Verbe ənel 'faire'<br>(arménien) | + | Flexion<br>Temps, personne, nombre<br>(arménien) |
|---|---|---|---|---|

Exemples (tirés de Kasparian 1992:292-383):

| | |
|---|---|
| E:1/239- "... vérifié *gəne*..." | (vérifié il fait) |
| S:1/416- "... engueulé *ənɨ*..." | (engueulé j'ai fait) |
| S:1/691- "... réglé *gəne*..." | (réglé il fait) |
| E:2a/40- "... varié *ənelu*..." | (varié pour faire) |
| E:2a/81- "... élaboré *ənem*..." | (élaboré je fais) |
| E:2a/77- "... observé *ənem*..." | (observé je fais) |
| S:2a/90- "... débranché *ənɨr en*..." | (débranché ils ont fait) |
| S:2a/154- "... refusé *gəne gor*..." | (refusé il est en train de faire) |

## 1. Pourquoi création et non emprunt ou calque?

### *1.a. Pourquoi est-ce une création et non un emprunt?*

La langue arménienne est une langue flexionnelle par excellence. Le verbe est formé de la racine (lexème) à laquelle est suffixé un morphème indiquant le temps, la personne et le nombre du verbe. Par exemple le verbe danser, *barel* donnera au temps du passé simple de l'indicatif (parfait):

- barec'i (j'ai dansé)
- barec'ir (tu as dansé)
- barec' (il a dansé)

Cette structure de formation de verbe favorise l'emprunt et l'intégration morphologique dans la construction de nouveaux verbes. Il suffit de remplacer le lexème du verbe en arménien par un lexème d'une autre langue et de garder la structure de conjugaison arménienne

(suffixes indiquant le temps, la personne et le nombre). Ce genre d'emprunts existe dans plusieurs situations de contact avec des langues flexionnelles comme dans les exemples de verbes anglais intégrés morphologiquement au français dans le québécois et surtout dans le parler chiac acadien: "j'ai watché ", "il a drivé", ou "il a kické". Nous avons plusieurs cas de formation de nouveaux verbes sur ce modèle dans la langue arménienne moderne parlée en Arménie ainsi qu'en diaspora.

Voici quelques exemples, de verbes formés à partir d'emprunts au russe, tirés d'une étude sur les emprunts et les calques dans la situation de contact de langues arménien-russe en Arménie ex-soviétique (cf. Dermerguerian 1984:55-56). Tel que le note l'auteur, dans ce cas de contact de langues, les formes verbales russes ne sont pas transférées en tant que telles, seuls les substantifs et quelques adjectifs russes empruntés permettent la formation de nouveaux verbes par l'ajout de suffixes flexionnels propres à l'arménien:

| substantifs russes | nouveaux verbes arméniens |
|---|---|
| montaž: "montage" | montaž**el**: "monter" |
| likvidac'ia: "liquidation" | likvidac'**nel**: "liquider" |
| efekt: "effet" | efekt**avorel**: "réaliser" |
| koordinac'ia: "coordination" | koordinac'**nel**: "coordonner" |
| **adjectif russe** | **nouveau verbe arménien** |
| potenc'ial: "potentiel" | potenc'**el**: "potentialiser" |

Dans ces cas, les emprunts sont intégrés morphologiquement et s'adaptent aux règles grammaticales qui régissent la structure interne des mots, propre à la langue arménienne (l'arménien forme la plupart de ses verbes à partir de noms ou d'adjectifs).

En ce qui concerne la situation de contact de langue au français, si le cas de construction verbale dont il est question est un

emprunt intégré à la structure arménienne, nous devrions avoir, suivant le même fonctionnement que les exemples cités ci-dessus, l'emprunt du substantif ou de l'adjectif ou de la racine du mot français intégré morphologiquement à la structure arménienne, ce qui nous donnerait pour le substantif "vérification", par exemple, la création d'un nouveau verbe "vérificationel", ou en considérant le participe passé du verbe français "vérifié" équivalent d'un adjectif, nous pourrions avoir la création du verbe "vérifiel" qui se conjuguerait au passé simple (parfait) de l'indicatif, de la manière suivante aux trois premières personnes du singulier:

| Verbe vérificationel | Verbe vérifiel |
|---|---|
| vérificationec'i | vérifiec'i |
| vérificationec'ir | vérifiec'ir |
| vérificationec' | vérifiec' |

Dans notre cas, la création ne suit aucun de ces modèles, car nous avons *vérifié* **ənel**: participe passé du verbe vérifier en français, conjugué avec le verbe faire en arménien.

*1.b. Pourquoi est-ce une création et non un calque d'une structure française?*

Si c'était le cas d'un emprunt lexical *vérifié* et du calque de la structure française du passé composé (verbe avoir conjugué au présent + passé du verbe conjugué) en arménien, nous devrions obtenir le verbe **unim** *vérifié* (j'ai vérifié) ou *vérifié* ***unim*** (vérifié j'ai). Ce n'est pas le cas puisque cette nouvelle structure verbale utilise le verbe **ənel** 'faire': *vérifié* **əri** (vérifié j'ai fait) et non l'auxiliaire avoir. Nous en concluons donc que cette nouvelle structure ne s'apparente ni à un calque ni à un emprunt du français en arménien, mais correspond plutôt à une nouvelle structure bilingue, créée grâce au contact des structures des deux langues arménienne et française pour répondre à des réalités et des nécessités de situations de communications bilingues, ces

nécessités pouvant être de plusieurs ordres : interactionnelle, sociolinguistique, psycholinguistique, etc.

Quels sont les *moyens linguistiques* à l'origine de ces créations? Nous allons essayer d'analyser les différents éléments qui constituent cette création et tenter d'émettre des hypothèses d'explication.

**2. Pourquoi le participe passé et non l'infinitif?**

Si nous prenons le premier constituant de cette création dans l'exemple "vérifié *ərɨ*", qui est un verbe français se terminant par le son /**e**/, nous pouvons nous poser la question à savoir pourquoi cela correspond au participe passé **é** et non à l'infinitif **er** du verbe français du 1$^{er}$ groupe.

Bien que la plupart de ces exemples de créations verbales concernent surtout les verbes français du 1$^{er}$ groupe, nous avons quelques cas de créations à partir de verbes français des 2$^{e}$ et 3$^{e}$ groupes qui confirment le temps des verbes français utilisé dans ces structures, c'est-à-dire le participe passé.

| | | |
|---|---|---|
| Ex: | finɨ *ərɨ* | (finɨ j'ai fait) |
| | battu *ərɨ* | (battu j'ai fait) |
| | mordu *ərɨ* | (mordu j'ai fait) |
| | déduit *ərɨ* | (déduit j'ai fait) |

Nous avons ainsi déterminé que la première partie de cette création correspond au participe passé du verbe français, la deuxième partie correspond au verbe *ənel* 'faire' conjugué à tous les temps et toutes les personnes des verbes de l'arménien. Sur quoi se base donc cette nouvelle création? Comment fonctionne-t-elle?

### 3. 1re hypothèse: *"Substantivation"* du participe passé du verbe français ou calque de la structure *pan mə ənel 'faire quelque chose'*

Notre première hypothèse présume que cette création fonctionne sur le modèle arménien *pan me ənel* 'faire quelque chose'. Dans ce modèle, le verbe faire est accompagné d'un substantif *pan* 'chose' qui se place devant le verbe ənel 'faire'. Dans les exemples suivants, *planté* əri (planté j'ai fait), *vérifié* əri (vérifié j'ai fait), le participe passé du verbe français, placé devant le verbe faire, jouerait le rôle d'un substantif. Nous aurions donc une "substantivation" du participe passé français dans la nouvelle structure en langue mixte. *Planté* fonctionnerait alors comme dans les exemples d'emprunts de substantifs français employés avec le verbe faire en arménien tel que:

E 1/404 " ...lutte **ənenk**... " (lutte faisons)
S 1/650 " ...frein **ənes**... " (frein tu fais)
S 1/685 " ...gueule **gəne**... " (gueule il fait)

Afin de confirmer cette première hypothèse, il faudrait que le participe passé "substantivé" vérifie les particularités du substantif arménien. Si nous appliquons les particularités du substantif arménien à l'exemple *vérifié* əri, nous nous apercevons que, dans cette structure, *vérifié* ne peut:

- prendre d'articles: *vérifié* ***n*** əri (**le** vérifié j'ai fait)
- avoir de marques du pluriel: *vérifié* ***ner*** əri (**des** vérifiés j'ai fait)
- être suffixé d'adjectif possessif: *vérifié* **s** əri (**mon** vérifié j'ai fait)
- avoir de désinences de cas, ex. l'instrumental: *vérifié ov* əri (**avec** vérifié j'ai fait)
- être précédé d'aucun adjectif: **ergu** *vérifié* əri (**deux** vérifiés j'ai fait)

L'adjectif placé devant *vérifié* ne le qualifierait pas mais qualifierait plutôt l'ensemble *"vérifié* ***əri*** ", c'est-à-dire le verbe, et aurait fonction

d'adverbe, exemple: ***ergar*** *vérifié* ***ərı*** **(long** vérifié j'ai fait)= (j'ai vérifié **longtemps)**.

Dans cette construction, *vérifié* ne suit donc aucune des règles de fonctionnement du substantif arménien[2]: il reste ***invariable*** et ***dépendant du verbe faire***. La première hypothèse de *"substantivation"* de l'élément participe passé du verbe français dans cette construction mixte, sur le modèle arménien, ***pan mə ənel*** 'faire quelque chose', n'est donc pas vérifiée. Ce qui nous amène à proposer une deuxième hypothèse.

**4. 2[e] hypothèse: "*Auxiliarisation*"[3] du verbe *ənel* 'faire'**

Dans cette nouvelle structure verbale, le participe passé français reste invariable et dépendant du verbe faire. Ses deux parties fonctionnent comme un amalgame. Le participe passé français nous donne le sens lexical du verbe, la suffixation de *ənel* 'faire' rajoute le sens de l'action, d'un mouvement dynamique et la flexion nous donne le temps, la personne et le nombre du verbe.

- planté ***ərı*** (j'ai fait)
- planté ***bidi ənem*** (je vais faire)
- planté ***gənem*** (je fais)

L'arménien utilise presque uniquement le verbe *ələl* 'être' comme auxiliaire.

[2] Pour une grammaire plus détaillée de la langue arménienne cf. Feydit (1969); Meillet (1962); Godel (1982).

[3] Cette hypothèse a été émise à la soutenance de thèse de doctorat de S. Kasparian par le membre du jury, Krikor Beledian, professeur de langue et de littérature arméniennes à l'Institut des Langues Orientales à Paris et à l'Université de Lyon.

- *hivant* ***em*** (malade je suis)
- *bargac* ***em*** (dormi je suis)

Nous pourrions aussi trouver, très rarement, dans la langue orale, ce même verbe dormir ainsi que d'autres verbes conjugués avec le verbe avoir ***unenal***, exemple:

- *bargac* ***unim*** (dormi j'ai)= j'ai envie de dormir
- *krac č'****unim*** (écris je n'ai pas)= je ne m'apprête pas à écrire

L'utilisation de l'auxiliaire avoir donne au verbe conjugué une dimension sémantique différente de l'auxiliaire être. L'auxiliaire *ǝlal* 'être' transmet un état passif de l'action (être dans un état de...) alors que l'auxiliaire *unenel* 'avoir' donne un sens de "vouloir être", "s'apprêter à être...". Le verbe *ǝnel* 'faire' n'est jamais utilisé comme un auxiliaire, le verbe *ǝlal* 'être' est le seul auxiliaire reconnu dans la langue arménienne.

Dans cette nouvelle création bilingue, le participe passé du verbe français est invariable, dépendant du verbe faire arménien et conjugué grâce à ce dernier verbe; ce qui confirme notre deuxième hypothèse. La création de ces nouveaux verbes bilingues utilise donc les moyens linguistiques suivants: *"auxiliarisation"* du verbe faire, *ǝnel* en arménien, et *amalgame* du participe passé du verbe français conjugué grâce à cet auxiliaire faire arménien. Les deux parties de cette création fonctionnent comme une unité sémantique et syntaxique.

D'autres cas de verbes bilingues composés avec le verbe *faire* sont relatés par Romaine (1995) dans la deuxième édition de son livre *Bilingualism*. Elle cite le cas des verbes composés bilingues, penjabi-anglais (cf. Emeneau 1956, Hook 1974), hindi-anglais (cf. Kachru 1978:36, Moag 1977, Bhatia 1982), tamoul-ânglais (cf. Annamalai 1978), japonais-anglais (cf. Stanlaw 1982, Nishimura 1986), turc-néerlandais (cf. Boeschoten et Verhoeven 1985:358), formés d'un nom\ adjectif\ verbe (dans une langue) + un opérateur ou l'auxiliaire faire

dans la langue réceptrice. D'après Romaine (1995:131), il semble que ces exemples ne constituent pas des exemples de restructuration, mais d'intégration dans un système réceptif. Ces modèles existent déjà dans le système de l'une des deux langues. Ces phénomènes de contact de langues sont considérés comme des alternances de langue ou des emprunts (il faudrait tenir compte de toute l'ambiguité qu'il y a autour de la terminologie et de la difficulté du classement des phénomènes de contact de langues).

Un seul cas de restructuration utilisant une construction périphrastique similaire entre l'espagnol et l'anglais est rapporté : le verbe faire conjugué en espagnol suivi d'un verbe anglais à l'infinitif. Il semble néanmoins que cette construction existe déjà, mais uniquement pour le causatif en espagnol et qu'elle soit seulement attestée en Californie (cf. Pfaff 1976:254-5). Dans le cas du modèle des verbes bilingues arméniens-français présentés dans cet article, la construction n'existe dans aucune des deux langues concernées.

Suite à ces études, il nous paraît intéressant de chercher à savoir si la construction de verbes mixtes avec le verbe faire est propre à toute étape de "créolisation", "pidginisation" ou "koinéisation" et, de façon plus générale, si le verbe **faire** n'a pas dans toute langue un statut qui pourrait le mener facilement dans une fonction d'auxiliaire.

Après avoir abordé ce cas dans une perspective surtout linguistique, nous aimerions conclure en abordant l'aspect pragmatique de ce mélange de langue. Pourquoi ces locuteurs bilingues utilisent-ils cette structure dans leur discours?

Nous notons que la création de ces nouveaux verbes bilingues n'a pas forcément pour fonction de décrire un concept nouveau (le mot arménien équivalent peut exister) ou de combler des lacunes du locuteur (le locuteur peut connaître le mot équivalent arménien). Elle n'est pas non plus uniquement le résultat d'une difficulté ou d'une facilité d'accès lexical du locuteur, une stratégie utilisée à des fins

interactionnelles, par exemple, une envie de se faire remarquer, de produire un effet comique ou de négocier son statut dans un groupe. Elle pourrait être le résultat de l'affirmation d'une identité bilingue à travers le choix de la création du terme dans la langue la plus appropriée à cette situation de communication bilingue qui serait la langue mixte. elle pourrait n'avoir pour fonction que la *créativité* (essence de la vie humaine) propre à l'individu bilingue, résultat du contact et de la fusion de deux langues, deux cultures, deux visions du monde...[4]

**Références bibliographiques**

Annamalai, E. 1978. "The anglicised Indian languages. A case of code-mixing". *International Journal of Dravidian Linguistics* 7:239-247.

Bhatia, T. 1982. "Trinidad Hindi. Three generations of transplanted language". *Studies in the linguistic Sciences* 11:135-51.

Boeschoten, H. et Verhoeven, L. 1985. "Integration niedelandischer lexicalischer Elemente". *Linguistische Berichte* 98:347-364.

Dermerguerian, Robert. 1984. *La langue arménienne en Arménie Soviétique, l'évolution du vocabulaire*. Publications Université de Provence, Marseille, France.

Emeneau Murray B. 1956. "India as a linguistic area". *Language* 32:3-16.

Feydit, Frederic. 1969. *Manuel de la langue arménienne, arménien occidental moderne*. Deuxième édition, Éditions Klincksieck, Paris.

Godel, Robert. 1982. *Linguistique arménienne, Études diachroniques*. Fondation des frères Goukassiantz - Vaduz.

[4] Remerciements à M. David Cohen, Professeur Émérite, Paris III, et École Pratique des Hautes Études Sorbonne, membre du CNRS Paris, (directeur de thèse); M. Krikor Beledian, Professeur de Langue et Littérature arméniennes à l'Institut des Langues Orientales, Paris et à l'Université de Lyon.

Hook, P. 1974. *The compound verb in Hindi*. The Michigan Series in South and Southwest Asian Linguistics I. Ann Arbor: University of Michigan.

Kachru, Braj B. 1978. "Toward structuring code-mixing, an Indian perspective". *International Journal of the sociology of language* 16:27-47.

Kasparian, Sylvia. 1992. *Analyse descriptive du "parler multilingue" d'Arméniens vivant à Paris, Approche interdisciplinaire*. Thèse de doctorat, Université de la Sorbonne Nouvelle PARIS III.

Meillet, Antoine. 1962. *Étude de linguistique et de philologie arméniennes, recherches sur la syntaxe comparée de l'arménien*. Bibliothèque Arménienne de la Fondation Calouste Gulbenkian, Lisbonne.

Moag, R. 1977. *Fiji Hindi*. Canberra: Australian National University Press.

Nishimura, M. 1986. "Intrasentential code-switching: the case of language assignment". In Vaid J. (réd). *Language Processing in Bilinguals: Psycholinguistic and Neuropsychological Perspectives*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 123-143.

Perrot, Marie-Ève. 1995. *Aspects fondamentaux du métissage français anglais dans le chiac de Moncton (N.-B.)*. Thèse de doctorat, Université de la Sorbonne Nouvelle, PARIS III (2 volumes).

Pfaff Carol W. 1976. "Functional and stuctural constraints on syntactic variation in code-switching". In Steever, S. B., Walker, C.A. and Mufwene, S.S. eds. *Papers from the Parasession on Diachronic Syntax*. Chicago: Chicago Linguistics Society, pp.248-259.

Romaine, Suzanne. 1995. *Bilingualism*. Blackwell Oxford UK et Cambridge USA.

Stanlaw J. 1982. "English in Japanese communicative strategies". In Kachru B. (réd.). *The Other Tongue. English Across Cultures*. Oxford: Pergamon Press, pp. 168-197.

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 201-213.

# Amérindianismes, variation formelle et dictionnaires

*Louis Mercier et Philippe Lanthier*
Université de Sherbrooke

## 1. Introduction

Au cours du XVII[e] siècle, le français a connu une période de contacts étroits avec les langues amérindiennes de la vallée du Saint-Laurent qui, correspondant pour les colons à une période d'adaptation au milieu nord-américain, était relativement favorable aux emprunts.[1] C'est à cette époque qu'ont été adoptés la plupart des mots d'origine amérindienne qui survivent en français québécois actuel. Ces amérindianismes sont relativement peu nombreux en comparaison des mots québécois d'origine anglaise. Cependant, en dépit de leur petit nombre, ils ne devraient pas être négligés de ceux qui s'intéressent à l'histoire de l'emprunt en français nord-américain. Par leur statut, nettement différent de celui des anglicismes, les amérindianismes constituent déjà une catégorie d'emprunts toute particulière. Ils ne sont pas perçus négativement. Bien au contraire, puisque certains d'entre eux, comme *achigan* « poisson d'eau douce », *atoca* « variété d'airelle » et *cacaoui* « canard marin », figurent dans la courte liste des premiers canadianismes de bon aloi reconnus par l'Office de la langue française (Beaulieu, Dupré et Canazzi 1969). Et leur intégration dans les dictionnaires français à partir des années 1970 n'a soulevé aucune controverse. Une autre particularité intéressante des amérindianismes tient à leur évolution formelle. Étant donné le caractère essentiellement oral des langues amérindiennes et la rapide diminution des contacts interlinguistiques, les mots amérindiens ont pénétré en français québécois sans tradition graphique et ont pu y évoluer sans vraiment subir de pressions externes, ce qui est loin d'être le cas des anglicismes.[2] Dans le présent article, il sera question de la variation formelle, tant phonétique que graphique, qui touche nos amérindianismes encore

---

1 Pour une description générale du phénomène de l'emprunt du français aux langues amérindiennes de l'est du Canada, v. Faribault 1993.

2 Comme les Québécois ont une bonne connaissance de la langue anglaise et qu'ils sont en contact continu avec cette langue, il est rare que nos anglicismes déclarés, même ceux qui sont bien intégrés au français québécois, échappent à toute influence formelle de leur étymon anglais. Il est peu probable, par exemple, que les graphies francisées *bine* et *lousse* parviennent un jour à éclipser totalement les graphies étymologiques *bean* et *loose*.

aujourd'hui et de la représentation qu'en donnent les dictionnaires généraux du français.

### 1.1 Variation formelle et représentation

Les enquêtes linguistiques effectuées au Québec et en Acadie au cours des dernières décennies nous ont permis de constater que, parmi les amérindianismes encore en usage, il y en avait fort peu à présenter une prononciation uniforme sur l'ensemble du territoire. Par exemple, les mots *babiche* « lanière de cuir non tanné » et *carcajou* « glouton » semblent appartenir à ce petit nombre. Cependant, dans la majorité des cas, incluant les trois premiers amérindianismes mentionnés, on observe une variation phonétique qui peut être plus ou moins importante.[3] Dans le cas du mot *achigan*, la variation se restreint à [aʃigɑ̃] ou [aʀʃigɑ̃] (PPQ q. 1410). Mais le mot *atoca* présente, quant à lui, toute une variété de prononciations dont les plus courantes sont [atakɑ], [atɔkɑ] et [atokɑ] (PPQ q. 1655). Les prononciations recueillies pour le mot *cacaoui*, sur lequel nous reviendrons plus loin, sont encore plus variées. Le constat de la variation phonétique qui touche encore aujourd'hui la plupart de nos amérindianismes nous a incités à nous intéresser à l'histoire de leur représentation graphique ainsi qu'à l'ensemble de leur représentation formelle dans les dictionnaires. Dans les pages qui suivent, nous ferons d'abord état, rapidement, de la variation graphique dont les dictionnaires actuels acceptent de rendre compte. Par la suite, nous étudierons l'ensemble des attestations phonétiques et graphiques disponibles du mot *cacaoui*, ce qui nous permettra finalement de comparer l'usage réel de ce mot et sa représentation lexicographique.

### 1.2 Le vocabulaire de la faune et de la flore

Un grande partie des amérindianismes qui sont encore en usage aujourd'hui relèvent du vocabulaire de la faune et de la flore.[4]

---

3 Nous ne pouvons aborder ici la question complexe des origines de cette variation.

4 Après avoir fait l'inventaire de tous les amérindianismes (présumés) signalés par ceux qui se sont intéressés à la question depuis la fin du XIX[e] siècle, Philippe Lanthier a dressé une liste d'environ 130 mots (sans compter les dénominations de peuples amérindiens) dont une cinquantaine réfèrent à des plantes ou des animaux. Grâce à l'appui financier du fonds FCAR et de

Ce sous-ensemble lexical, qui fonctionne indépendamment du thème amérindien, est particulièrement bien intégré au lexique québécois.

Quelques-uns de nos amérindianismes usuels jouissent d'une longue tradition dans la lexicographie française: c'est le cas de *caribou* « grand cervidé nordique », *carcajou* et *pékan* « grande martre ».[5] Mais la plupart d'entre eux ont dû attendre le *Dictionnaire général de la langue française au Canada* (1957) de L.-A. Bélisle avant de pénétrer la nomenclature des dictionnaires généraux. Nous avons réuni dans le tableau n° 1 tous les noms d'animaux et de plantes généralement reconnus comme emprunts directs aux langues amérindiennes de l'est du Canada qui figurent dans l'un ou l'autre des dictionnaires généraux du français les plus largement diffusés sur le marché québécois depuis la parution du premier dictionnaire général québécois.[6] Il s'agit, du côté québécois, du dictionnaire de Bélisle (Bélisle 1957-1979), du *Dictionnaire du français Plus* (DFP 1988), du *Dictionnaire québécois d'aujourd'hui* (DQA 1992) et, du côté français, des *Petit Larousse* (PL 1972-1994) et *Petit Robert* (PR 1977-1993). À ces sources, nous ajoutons les *Canadianismes de bon aloi* de l'Office de la langue française (OLF 1969) dont il a été question plus tôt. Notre tableau fait abstraction des prononciations proposées pour ne faire état que des graphies présentées en vedette ou signalées dans le corps d'un article.[7]

Tous les mots québécois d'origine amérindienne qui sont relativement bien connus ont déjà été répertoriés par au moins un

---

l'Université de Sherbrooke, le vocabulaire québécois de la faune et de la flore fait l'objet de divers travaux de nature lexicographique au Centre d'analyse et de traitement informatique du français québécois de l'Université de Sherbrooke.

5 C'est également le cas de *ondatra*, synonyme technique de *rat musqué*, mais cet amérindianisme, dont on attribue l'emprunt au naturaliste français Buffon, ne semble jamais avoir eu cours dans la langue générale (Faribault 1993 : 200-201).

6 Nous ne retenons pas les mots comme *wapiti* qui sont d'origine amérindienne mais nous sont très probablement parvenus par le biais de l'anglais.

7 Pour faciliter la lecture du tableau n° 1 : V = vedette unique; 1 ou 2 = première ou deuxième forme d'une vedette double; R = vedette d'un article ne contenant qu'un renvoi; C = forme citée dans le corps d'un article consacré à une autre forme. Arbres et plantes: *atoca, chicoutée, mascouabina, pacane, pénac, pimbina, saccacomi, savoyane*. Mammifères: *caribou, carcajou, ondatra, pékan, pichou*. Oiseaux: *cacaoui, moyac*. Batracien: *ouaouaron*. Poissons: *achigan, malachigan, maskinongé, ouananiche, ouitouche, poulamon, touladi*.

| Tableau n° 1 | Bélisle 1957-1979 | OLF 1969 | DFP 1988 | DQA 1992 | PL 1972 -> | PR 1977 -> |
|---|---|---|---|---|---|---|
| achigan | V | V | V | V | V | V |
| ataca | 2R *atacas* | - | 2R | 2R | 80: - > 91: 2R | 77: 2 > 93: CR |
| atoca | 1 *atocas* | V | 1 | 1 | 80: V > 91: 1 | 77: 1 > 93: V |
| cacaoui | - | V | 1 | 1 | 1991 -> | V |
| kakawi | - | - | 2R | 2R | - | - |
| caribou | V | - | V | V | V av. 1950 -> | V 1967 -> |
| carcajou | V | - | V | V | V av. 1950 -> | V 1967 -> |
| chicotée | 2 | - | - | - | - | - |
| chicouté | 1 | - | - | - | - | - |
| chicoutée | C *plaqueb.* | - | - | - | - | - |
| malachigan | V | - | V | V | - | - |
| mascou | V | | 1 | 2 | - | - |
| mascabina | 1 | - | - | - | - | - |
| mascouabina | - | - | 2R | 1 | - | - |
| mascoubina | C *sorbier* | - | - | - | - | - |
| maskouabina | 2 | - | - | - | - | - |
| maskinongé | V | V | V | V | V | V |
| moïac | - | - | - | 2R | - | - |
| moignac | C *eider* | - | - | - | - | - |
| moyac | - | - | - | 1 | - | - |
| ondatra | V | - | V | V | V av. 1950 -> | V 1967 -> |
| ouananiche | V | V | V | V | V | V |
| ouaouaron | V | V | - | V | V | V |
| wawaron | R | - | - | - | - | - |
| ouitouche | 1979 | - | - | - | - | - |
| pacane | V | - | V | V | V 1991 -> | V |
| pécan | R *pacane* | - | V | C | - | 1 |
| pecan | - | - | - | - | - | 2 |
| pécane | V | - | - | - | - | - |
| pécan | 2R | - | 2R | 2R | - | - |
| pécant | R | - | - | - | - | - |
| pékan | 1 | - | 1 | 1 | V av. 1950 -> | V 1967 -> |
| pénac | V *pénacs* | - | - | - | - | - |
| pichou | V «soulier» | - | V | V | - | - |
| pembina | - | - | 2R | 2R | - | 77: 2 > 93: CR |
| pimbina | V | - | 1 | 1 | V 1991 -> | 77: 1 > 93: V |
| poulamon | V | - | V | V | - | - |
| saccacomi | - | - | - | 1 | - | - |
| saccacomie | - | - | - | 2 | - | - |
| saouiyane | 2 | - | - | - | - | - |
| savoyane | 1 | - | - | V | - | - |
| touladi | V | - | V | V | - | - |
| touradi | - | - | - | C | - | - |

dictionnaire québécois et figurent dans notre tableau. On ne sera pas étonné de constater que les nomenclatures québécoises s'ouvrent davantage à nos amérindianismes — le dictionnaire québécois le plus récent en relève deux fois plus que les dernières éditions du PL et du PR — et rendent davantage compte de leur variation.

Le plus grand intérêt de notre tableau est de mettre en évidence les différences de traitement qui existent entre les trois dictionnaires québécois et plus particulièrement entre le Bélisle et le DQA. Si l'on s'en tient uniquement aux variantes, on fera les constatations suivantes. Certaines formes mentionnées par Bélisle sont passées sous silence dans le DQA (*mascabina*, *wawaron*, *saouiyane*) ou délaissées au profit d'autres formes (*moignac* > *moïac*, *moyac*; *maskouabina* > *mascouabina*). Par ailleurs, le DQA signale des cas de variation ignorés de Bélisle (*pembina / pimbina*, *touladi / touradi*). On peut se demander si les modifications récemment apportées à la représentation lexicographique correspondent vraiment à des changements dans l'usage. D'autres modifications seraient-elles souhaitables? Pour être en mesure de bien évaluer la représentation lexicographique de nos emprunts aux langues amérindiennes, nous avons entrepris de nous renseigner sur leur usage réel. L'étude qui suit est révélatrice du type de compléments et de correctifs qui peuvent encore être apportés à la description lexicographique pour la rendre plus fidèle à l'usage.

## 2. Étude du mot *cacaoui*[8]

L'amérindianisme *cacaoui*, qui est attesté en français nord-américain depuis le XVII[e] siècle, semble d'origine algonquienne et avoir été emprunté au micmac (DFQ). Le mot sert à désigner un canard marin (*Clangula hyemalis*) qui passe l'été dans les régions arctiques (Amérique, Europe, Asie) mais se tient plus au sud en hiver, saison pendant laquelle il est présent sur tout le pourtour de l'estuaire et du golfe du Saint-Laurent (Godfrey 1986 : 122). Cette espèce de canard n'est pas aussi largement connue du grand public que les canards noir ou malard (colvert) mais, d'après les enquêtes de Gaston Dulong (PPQ q. 1489; v. aussi q. 219 et q. 1737) et de Geneviève Massignon (q. 435), tous ceux qui la connaissent au Québec et en Acadie la désignent spontanément sous le nom de

[8] Nous adoptons ici la graphie qui est la plus largement répertoriée dans les dictionnaires actuels du français.

*cacaoui*.[9] Ce mot est aussi en usage chez les spécialistes québécois qui, depuis 1957 (SCF : 7), l'ont adopté comme élément spécifique, sous la forme *kakawi*, pour créer la dénomination bi-nominale officielle *canard kakawi*, récemment modifiée en *harelde kakawi* .

### 2.1 Ce qu'en disent les dictionnaires

Comme le montre le tableau n° 1, le mot *cacaoui* n'est entré dans les dictionnaires généraux du français qu'à la fin des années 1970 (PR 1977). Depuis cette date, tous les dictionnaires ne donnent qu'une seule prononciation : [kakawi]. Pour ce qui est de la graphie, les dictionnaires de France (PR et PL) ne mentionnent que *cacaoui*. Seuls les dictionnaires québécois (DFP et DQA) font état d'une variation graphique; mais, s'ils répertorient en outre la forme *kakawi*, ils n'apportent toutefois sur cette variation aucune précision qui puisse guider leurs lecteurs.

### 2.2 Ce qu'il en est de l'usage

*Les attestations phonétiques* — Le tableau suivant présente, en API, tout ce que nous avons pu réunir comme attestations phonétiques du mot *cacaoui*. Le plus grand nombre de ces attestations (38 sur 43) provient de l'atlas de Dulong.

| Tableau n° 2 | 15 | [kakawi] | 2 | [kakawit] | 1 | [kakawik] |
|---|---|---|---|---|---|---|
| | 6 | [kakawi] | 2 | [kākawit] | 1 | [kakawiz] |
| | 6 | [kakaɥit] | 1 | [kakawi] | 1 | [kakawɛt] |
| | 4 | [kakawi] | 1 | [kakawit] | 1 | [kakwi] |
| | | | 1 | [kakawit] | | |

Les données phonétiques disponibles sont donc peu nombreuses et relativement anciennes.[10] Néanmoins, elles suffisent

9 Le mot *cacaoui* ne figure pas parmi les données recueillies lors de l'enquête de Thomas Lavoie (Lavoie, Bergeron et Côté 1985, q. 716); mais, d'après les réponses de ses témoins, trop vagues (*canard, canard sauvage*) ou trop hétéroclites (*canard blanc, caille, gris, noir*, etc.), on peut penser que ceux-ci n'étaient pas familiers avec le référent.

10 Dans les îles françaises voisines de Terre-Neuve, Brasseur et Chauveau ont relevé [kakawi] à St-Pierre, [kakawit] et [kakawik] à Miquelon (1990 : 142).

pour confirmer l'existence d'une variation et dessiner les grandes lignes de cette variation.

*Les attestations graphiques* — Toutes les attestations graphiques du mot *cacaoui* que nous avons pu réunir sont présentées dans le tableau n° 3. Les 17 formes recueillies sont classées d'après la date de leur première attestation; lorsqu'elles sont attestées plus d'une fois, nous précisons également la date de leur dernière attestation. Nous donnons en outre le nombre de sources où ces formes ont été relevées. Pour dresser cette liste, nous avons consulté le fichier linguistique et l'*Index lexicologique québécois* du Trésor de la langue française au Québec (Poirier et Mercier 1986), la banque de données textuelles du Centre d'analyse et de traitement informatique du français québécois de l'Université de Sherbrooke, ainsi que la banque d'articles de presse québécois disponibles sur disques optiques (*Actualité/Québec* 1993-1995). Les données graphiques dont nous disposons sont nettement plus nombreuses que les données phonétiques. Elles s'étalent sur une plus grande période.

Tableau n° 3

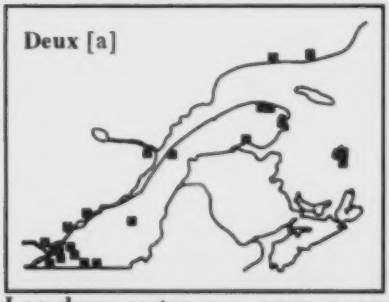
| | | | | | |
|---|---|---|---|---|---|
| cacaoüy | 1 | 1672 | cacowi | 1 | 1902 |
| **cacaoui** | 22 | 1760 —> | cacawit | 1 | 1909 |
| cancanwi | 2 | 1876-1888 | kacaoui | 1 | 1916 |
| kakaoui | 3 | 1885-1928 | kacarlic | 1 | 1916 |
| **kakawi** | 32 | 1883 —> | kacauouic | 1 | 1916 |
| takaoui | 1 | 1889 | cacaouit | 2 | 1925-1927 |
| cacaouite | 2 | 1894-1989 | cacaouït | 1 | 1928 |
| cacaouïte | 1 | 1897 | kacaouite | 1 | 1951 |
| cacawi | 1 | 1900 | | | |

## 2.3 Variation phonétique et correspondance graphique

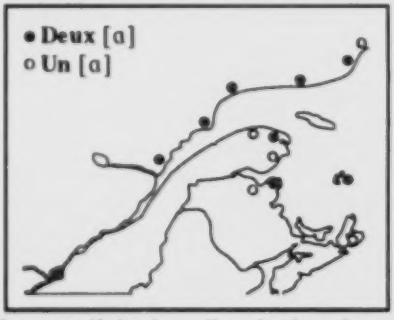
Du côté des prononciations, la forme [kakawi], qui est la seule à être signalée par les dictionnaires français et québécois, se présente clairement comme la forme dominante avec plus du tiers des attestations, qui sont réparties sur l'ensemble du territoire québécois. Viennent ensuite 3 prononciations secondaires, attestées de 4 à 6 fois, et une douzaine d'autres variantes phonétiques marginales attestées une fois ou deux. Du côté des graphies, 2 formes se dégagent nettement : *cacaoui* et *kakawi*. Les autres variantes sont très marginales et complètement absentes de la documentation récente. On

observe une forte concurrence entre *cacaoui* et *kakawi*, mais cette concurrence est essentiellement graphique puisque ces variantes peuvent indifféremment être associées à la prononciation dominante, tout comme au moins 4 autres variantes graphiques (*cacaoüy*, *kakaoui*, *cacawi* et *kacaoui*). Nous reviendrons sur cette concurrence graphique, mais nous passerons d'abord en revue les principaux cas de variation phonétique. Prises indépendamment, les 3 prononciations secondaires ([kakawi], [kakawi] et [kakaɥit]) peuvent sembler négligeables aux lexicographes. Mais lorsqu'on les compare à l'ensemble des variantes phonétiques et graphiques relevées, on constate qu'elles illustrent des éléments de variation dont certains sont assez largement répandus pour mériter d'être signalés.

*Le timbre du /A/* — Les formes [kakawi] et [kakawi] se distinguent de la prononciation dominante [kakawi] par la présence d'au moins un [ɑ] postérieur. Ce trait n'est pas négligeable puisqu'il se retrouve dans plus du tiers des prononciations recueillies (7 fois dans les deux syllabes; 6 dans la seconde syllabe et 2 dans la première syllabe).

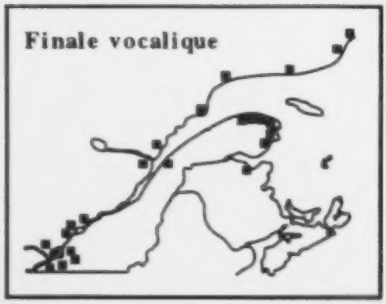



Les deux cartes que nous avons réalisées à partir des données du PPQ, nous montrent clairement que, contrairement aux prononciations avec 2 [a] antérieurs, les prononciations avec au moins 1 [ɑ] postérieur sont toutes localisées le long de l'estuaire et du golfe du Saint-Laurent, régions où le canard en question est sans doute mieux connu. Par rapport au territoire québécois, il s'agit donc d'importantes prononciations régionales. Dans le domaine acadien (Provinces maritimes, Baie-des-Chaleurs, Îles-de-la-Madeleine et Côte-Nord), il s'agirait même des seules prononciations largement répandues. La graphie marginale *cacowi*, attestée sous la plume de J. Geddes, philologue américain familier avec le parler de la Baie-



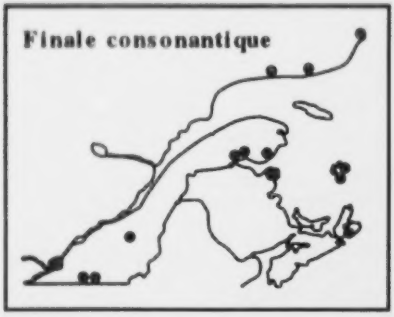
des-Chaleurs, peut sans doute être interprétée comme une tentative de transcription de la prononciation [kakɑwi].

*La finale consonantique* — La forme [kakaqit] se distingue de la prononciation dominante à la fois par sa semi-consonne antérieure [ɥ] et sa finale consonantique. C'est la seule de toutes les prononciations relevées à présenter un [ɥ] plutôt qu'un [w]. Il s'agit donc d'un trait relativement marginal. 4 attestations sur 6 de cet emploi ont été relevées autour du Golfe, mais sa faible fréquence ne permet pas de le dégager comme trait régional. Du côté des graphies, aucune forme ne peut lui être associée. En revanche, [kakaqit] est apparentée par sa finale consonantique à 4 autres variantes phonétiques terminées en [t] ([kakawit], [kɑkɑwit], [kakɑwit] et ([kakawɛt]), et même à 5 si l'on ajoute celle qui se termine en [k] ([kɑkɑwik]).




Ce trait est présent dans le tiers de nos attestations phonétiques; il ne s'agit donc pas d'un trait négligeable. Du côté des graphies marginales, il est assez clair que les formes *cacaouite*, *cacacoüite* et *kacaouite* veulent rendre compte d'une finale consonantique; on peut penser que c'est également le cas de *cacawit*, *cacaouit*, *cacaoüit*, *kacauouic* et *kacarlic*.

Selon les données du PPQ, la finale consonantique est nettement plus fréquente dans l'est du Québec; elle serait même dominante dans la zone acadienne.[11] Au cours de leur enquête sur le parler des îles françaises de Saint-Pierre et de Miquelon, qui présente certaines affinités avec le parler acadien,



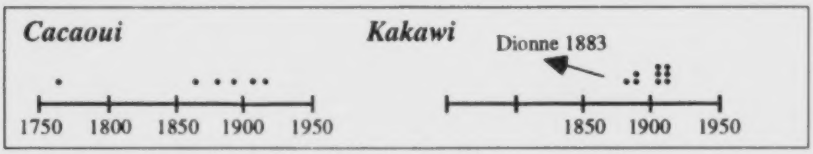

11 Ce qui donne à penser que les formes *cacaouit*, *cacaoüit*, *kacauouic* et *kacarlic*, relevées sous la plume de l'homme de lettres acadien Pascal Poirier, voulaient vraisemblablement rendre compte d'un finale consonantique.

P. Brasseur et J.-P. Chauveau (1990 : 142) n'ont d'ailleurs relevé à Miquelon que des formes à finale consonantique ([kakawit] et [kakawik]).

*Variation phonétique et dictionnaires* — Comme nous avons pu le constater, la prononciation proposée par les dictionnaires du français correspond à celle qui est la plus fréquente et la seule d'emploi généralisé. Mais nous avons vu également que cette prononciation était fortement concurrencée, dans l'est du territoire québécois et en Acadie, par des variantes avec [ɑ] postérieur ou finale consonantique en [t]. Il y aurait sans doute lieu de mentionner ces variantes dans un dictionnaire soucieux de rendre compte des parlers régionaux québécois; elles devraient nécessairement figurer dans un dictionnaire du parler acadien.[12]

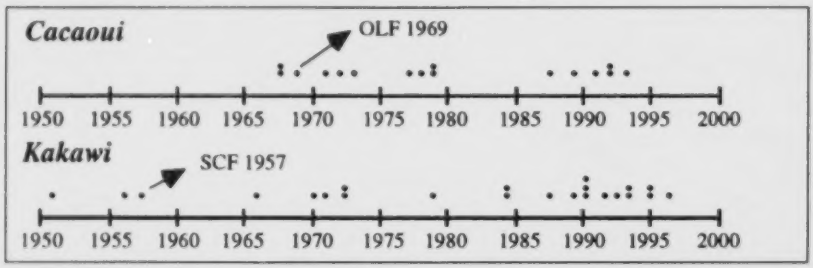
## 2.4 Variation essentiellement graphique

Quant aux deux seules variantes graphiques importantes du mot à l'étude, soit *cacaoui* et *kakawi*, elles sont déjà mentionnées dans les deux derniers dictionnaires généraux québécois parus (DFP et DQA). Mais le lecteur de ces dictionnaires ne trouve aucune information lui permettant de situer ces variantes l'une par rapport à l'autre, ni d'orienter son choix. Avant de comparer l'histoire de ces deux formes dominantes, notons qu'elles s'opposent dans leur représentation des sons [k] et [w]. Les graphèmes *c* et *ou* de *cacaoui* correspondent davantage à la tradition française alors que les graphèmes *k* et *w* de *kakawi* trahissent l'origine étrangère du mot (Catach 1988 : 148-149; Yaguello 1990 : 41-42). Par leur association de graphèmes, ces graphies sont cohérentes contrairement à bon nombre de graphies marginales (*kakaoui*, *cacawi*, *kacaoui*, etc.).




12 La vitalité actuelle de ces variantes resterait cependant à être vérifiée.

Comme on le voit ci-dessus, *cacaoui* remonte au milieu du XVIII[e] s., mais ses attestations seront relativement rares au cours des deux siècles suivants. [13] Celles qui ont été recueillies autour de 1900 sont exclusivement dues à des philologues (Chamberlain 1888 : 259; Clapin 1894 : 349, Geddes 1908 : 260; Poirier 1916 : 344). La graphie *kakawi* n'est apparue qu'à la fin du XIX[e] s., sous la plume de l'ornithologue C.-E. Dionne (1883 : 224). Dans les années suivantes, elle réapparaîtra aussi bien sous la plume des naturalistes (Puyjalon 1893 : 178, Provancher 1904 : 173; Huard 1912 : 581) que celle des philologues (Chamberlain 1894 : 43; Dionne 1909 : 404; Poirier 1916 : 345). La dernière figure nous donne un aperçu de l'étalement et de la fréquence des deux variantes après 1950, période d'où proviennent la majorité de leurs attestations.




Absente de la documentation pendant près d'un demi-siècle, la graphie *cacaoui* a été relancée à la fin des années 1960 quand le mot a été retenu dans la liste des canadianismes de bon aloi préparée par l'Office de la langue française (OLF 1969). Par la suite, on la retrouve presque exclusivement dans des répertoires lexicologiques et lexicographiques. Depuis Dionne (1883), les naturalistes et ornithologues n'ont jamais cessé de lui préférer la forme *kakawi*. C'est d'ailleurs sous cette forme qu'en 1957, le mot est officiellement entré comme terme spécifique dans la terminologie ornithologique québécoise (SCF 1957 : *canard kakawi*). Depuis le début des années 1980, les attestations de *kakawi* sont plus fréquentes que celles de *cacaoui*, et cela même dans les sources non spécialisées. Dans les

13 Chacun des points correspond à une source d'attestation. *Cacaoui* figure dans un document acadien de 1758; la première variante attestée (*cacaotty* 1672) provient également d'un document d'origine acadienne (Massignon 1962 : 275).

grands périodiques québécois parus au cours des trois dernières années (*Actualité/Québec*), nous n'avons relevé que la forme *kakawi*.

*Variation graphique et dictionnaires* — La graphie *cacaoui*, qui continue à être portée par la tradition lexicographique, montre des signes évidents de vieillissement dans l'usage réel. Seule la graphie *kakawi* semble actuellement en usage et la situation ne risque guère d'être modifiée prochainement étant donné le prestige dont jouit la langue spécialisée, de plus en plus largement diffusée. Depuis 1993, *kakawi* est entré dans la terminologie ornithologique unifiée des pays francophones (*harelde kakawi*; CINFO : 14). Aucun dictionnaire du français ne devrait dorénavant faire l'économie de cette graphie.

Comme on peut le constater, les dictionnaires actuels du français ne donnent pas une image tout à fait fidèle de la variation formelle qui continue, encore aujourd'hui, à affecter certains de nos amérindianismes. Il y a donc lieu de travailler à parfaire cette réprésentation.

## BIBLIOGRAPHIE

*Actualité / Québec*. 1993-1995. Société nationale d'information.

Beaulieu, Maurice, Céline Dupré, & Jean-Baptiste Canazzi. 1969. *Canadianismes de bon aloi* (OLF). (Cahiers de l'Office de la langue française, n° 4). Québec : Ministère des Affaires culturelles.

Bélisle, Louis-Alexandre. 1957. *Dictionnaire général de la langue française au Canada*, Québec : Bélisle éditeur. 3e éd. 1979. *Dictionnaire nord-américain de la langue française*. Montréal : Beauchemin.

Brasseur, Patrice & Jean-Paul Chauveau. 1990. *Dictionnaire des régionalismes de Saint-Pierre et Miquelon*, Tübingen : Niemeyer.

Catach, Nina. 1986. *L'orthographe française : traité théorique et pratique*. Paris : Nathan.

Chamberlain, A. F. 1888. « Indiano-Canadian Words III ». *American Notes and Queries* 1, 29 sept. : 258-259.

Chamberlain, A. F. 1894. « The Life and Growth of Words in the French Dialect of Canada ». *Modern Language Notes* 9,2 : 39-44.

Clapin, Sylva. 1894. *Dictionnaire canadien-français*. Montréal—Boston : Beauchemin & Fils—Sylva Clapin (réimpr. Québec : Les Presses de l'Université Laval. 1974).

Commission internationale des noms français des oiseaux (CINFO). 1993. *Les noms français des oiseaux du monde*. Sainte-Foy—Bayonne : MultiMondes—Chabaud.

*Dictionnaire du français Plus* (DFP). 1988. Montréal : CEC.

*Dictionnaire du français québécois* (DFQ). À paraître. Trésor de la langue française au Québec (Université Laval).

*Dictionnaire québécois d'aujourd'hui* (DQA). 1992. DicoRobert.

Dionne, Charles-Eusèbe. 1883. *Les oiseaux du Canada*. Québec : P.-G. Delisle.

Dionne, Narcisse-Eutrope. 1909. *Le parler populaire des Canadiens français*. Québec : Laflamme et Proulx imprimeurs (réimpr. Québec : Les Presses de l'Université Laval. 1974).

Dulong, Gaston & Gaston Bergeron. 1980. *Le parler populaire du Québec et de ses régions voisines* (PPQ). Québec : Ministère des Communications du Québec et Office de la langue française.

Faribault, Marthe. 1993. « L'emprunt aux langues amérindiennes ». *Français du Canada - français de France. Actes du troisième Colloque international d'Augsbourg du 13 au 17 mai 1991*. Tübingen : Niemeyer, p. 199-215.

Geddes, James Jr. 1908. *Study of an Acadian-French Dialect Spoken on the North Shore of the Baie-des-Chaleurs*. Halle : Max Niemeyer Publisher.

Godfrey, W. Earl. 1986. *Les oiseaux du Canada*. Ottawa : Musée national des sciences naturelles—Musées nationaux du Canada.

Huard, Victor-Alphonse. 1914. « La terminologie franco-canadienne dans les sciences naturelles ». *Premier Congrès de la langue française au Canada 1912. Mémoires*. Québec : Imprimerie de l'Action nationale limitée, pp. 573-588.

Lavoie, Thomas, Gaston Bergeron et Michelle Côté. 1985. *Les parlers français de Charlevoix, du Saguenay, du Lac-Saint-Jean et de la Côte-Nord*. Montréal : Les Publications du Québec.

Massignon, Geneviève. [1962]. *Les parlers français d'Acadie. Enquête linguistique*. Paris : Klincksieck.

Poirier, Claude & Louis Mercier (dir.). 1986. *Index lexicologique québécois*. Université Laval (Québec) : Trésor de la langue française au Québec.

Poirier, P. 1916. « Des vocables algonquins, caraïbes, etc., qui sont entrés dans la langue ». *Mémoires et comptes rendus de la Société royale du Canada* 10,1, section 1 : 339-364.

Provancher, Léon. 1904 janv. *Le naturaliste canadien*. 31, 1.

Puyjalon, Henri de. 1893. *Guide du chasseur de pelleterie*. Montréal : P. J. Bedard.

Service canadien de la faune sauvage (SCF). 1957. *Liste des noms français des oiseaux du Canada*. Ottawa.

Yaguello, Marina. 1990. *Histoire de lettres*. Paris : Seuil.

Julie Auger et Yvan Rose (réd.). (1997). *Explorations du lexique*. Québec: CIRAL, 215-228.

# Préservation phonémique et seuil de tolérance dans l'adaptation d'emprunts

*Carole Paradis*
Université Laval

## 1. INTRODUCTION

Cet article veut montrer que l'adaptation phonologique d'emprunts ne se fait pas de façon arbitraire. En effet, on peut prédire dans une large mesure le type d'adaptation que subissent les segments étrangers ainsi que les rares cas où ils sont élidés, à partir des principes phonologiques de la *Théorie des contraintes et des stratégies de réparation* (TCSR; cf. Paradis 1988a,b, et Paradis et LaCharité 1993).

Un emprunt peut être défini comme un mot (simple ou composé) de la langue source qui est incorporé dans le discours de la langue emprunteuse et qui respecte au moins les contraintes les plus périphériques de cette dernière (cf. Paradis et Lebel, ce volume, pour un exposé sur la périphérie). Sur la base de nombreuses études sociolinguistiques (Haugen 1950, Mougeon, Beniak et Valois 1985 et Poplack, Sankoff et Miller 1988) et phonologiques (Paradis et LaCharité 1996, 1997), il est tenu pour acquis que les emprunts sont introduits par des «bilingues»[1], qui, par définition, ont accès au code phonologique de la langue source. Les emprunts sont introduits à travers ce que les sociolinguistes appellent des *code-switches*, des *nonces* et des idiosyncrasies. L'étude de Haugen (1950), en particulier, révèle que non seulement les emprunts sont-ils introduits par les bilingues mais que les patrons phonologiques d'adaptation sont aussi imposés par les bilingues; les patrons d'adaptation prennent le statut de conventions sociales au fur et à mesure que le bilinguisme s'installe dans la communauté emprunteuse.

Le reste de l'article est divisé en trois parties. En 2, sont présentés les statistiques tirées de nos corpus d'emprunts ainsi que les principaux mécanismes de la TCSR. Nous verrons en 3 comment certains segments et structures du français sont adaptés en peul et, en 4, pourquoi certains de ces segments sont élidés plutôt qu'adaptés. Enfin, une conclusion est offerte en 5.

---

[1] Le terme doit être pris dans son sens large; il fait référence à tous les types de bilingues, non seulement aux bilingues dits «équilibrés», c'est-à-dire aux bilingues aussi compétents dans la langue A que dans la langue B.

## 2. STATISTIQUES ET TCSR

Nos études sur les emprunts sont basées sur plusieurs corpus. Ceux-ci constituent, à ma connaissance, les plus gros corpus d'emprunts au monde transcrits en alphabet phonétique international et étudiés d'un point de vue phonologique. Le tableau (1) présente les statistiques tirées de quatre corpus, soit celui d'emprunts anglais en français québécois de Québec et ceux d'emprunts français en arabe marocain, en kinyarwanda et en peul. Dans la colonne de droite du tableau, on peut observer que les quatre corpus comprennent 3796 emprunts et 9103 formes. Une «forme» est un emprunt prononcé par un informateur, c'est-à-dire la réalisation concrète d'une entité abstraite, l'emprunt. Cette distinction entre «forme» et «emprunt» est essentielle puisque les informateurs ne prononcent pas tous les emprunts de la même manière, d'une part, et qu'un emprunt n'est pas nécessairement connu de tous les informateurs, d'autre part. Comme chaque emprunt a été présenté à trois locuteurs natifs de la langue emprunteuse, un même emprunt, s'il est connu des trois locuteurs, peut générer jusqu'à trois formes. Les formes contiennent 12438 malformations phonologiques, ce qui nous permet de dégager des tendances très nettes quant aux patrons d'adaptation et d'élision.

(1)

| nombre de: | Emprunts anglais en *français québécois* | Emprunts français en: *arabe marocain* | *kinyarwanda* | *peul* | Total |
|---|---|---|---|---|---|
| emprunts | 948 | 1547 | 756 | 545 | 3796 |
| formes | 2412 | 3512 | 2143 | 1036 | 9103 |
| malformations | 3501 | 3660 | 4444 | 833 | 12438 |
| adaptations | 2521 (/3501 =72,0%) | 2981 (/3660 =81,4%) | 4313 (/4444 =97,1%) | 785 (/833 =94,2%) | 10600 (/12438 =85,2%) |
| élisions | 90 (/3501 =2,6%) | 230 (/3661 =6,3%) | 104 (/4444 =2,3%) | 32 (/833 =3,9%) | 456 (/12438 =3,7%) |
| non-adaptations | 885 (/3501 =25,4%) | 449 (/3661 =12,3%) | 27 (/4444 =0,6%) | 16 (/833 =1,9%) | 1377 (/12438 =11,1%) |

adaptation = substitution segmentale ou insertion d'un segment; élision = élision d'un segment; non-adaptation = structure ou segment étranger laissé tel quel dans la langue emprunteuse.

Comme indiqué dans la colonne de droite, les élisions sont rares (3,7% des cas). Le fort taux de préservation segmentale est attribué au *principe de préservation* en (2).

(2) *Principe de préservation*:
L'information segmentale est préservée maximalement, dans les limites du seuil de tolérance.

Un segment mal formé n'est élidé que lorsque sa préservation est trop coûteuse. La limite du «coût» qu'une langue peut se permettre pour préserver un segment est établie par le *principe du seuil de tolérance* en (3). Ce principe pose que toutes les langues imposent une limite à la préservation segmentale (la partie la plus importante du principe) et que cette limite consiste en un certain nombre d'étapes ou de réparations permises à l'intérieur d'un domaine donné.

(3) *Principe du seuil de tolérance*:
a) toutes les langues établissent une limite à la préservation segmentale;
b) cette limite est de *x* étapes à l'intérieur d'un domaine de contrainte donné.[2]

Le non-respect du seuil de tolérance entraîne une élision segmentale. Autrement, un segment mal formé est substitué à un autre ou une structure non syllabable réparée par épenthèse (85,2% des cas). Dans le cadre adopté ici, les réparations (substitutions, épenthèses et élisions) sont effectuées par des *stratégies de réparation*.

(4) *Stratégie de réparation*: une opération phonologique universelle et non contextuelle dont l'application est causée par la violation d'une contrainte phonologique et qui insère ou élide du matériel phonologique dans le but de satisfaire la contrainte violée.

Les stratégies de réparation s'appliquent minimalement.

(5) *Principe de minimalité*: les stratégies de réparation s'appliquent
a) au niveau le plus bas auquel fait référence la contrainte violée
b) en impliquant le moins d'étapes (d'opérations) possibles.

Le niveau phonologique le plus bas est déterminé par la *hiérarchie des niveaux phonologiques* en (6).

(6) *Hiérarchie des niveaux phonologiques (HNP)*: niveau métrique > niveau syllabique > niveau du squelette > noeud de base > traits non terminaux > traits terminaux.

---

[2] La limite observée en peul et en kinyarwanda est de deux étapes. Toutefois, il semblerait qu'elle puisse varier d'une langue à l'autre. Par exemple, dans le corpus d'emprunts français en lingala que nous sommes en train de construire, la limite est plus élevée, ce qui conduit à un plus petit pourcentage d'élisions.

La HNP permet aussi de déterminer quel est le niveau le plus élevé dans la convention (7).

(7) *Convention de préséance*: dans le cas où deux ou plusieurs contraintes sont violées, la contrainte qui a priorité est celle qui fait référence au niveau le plus élevé dans la HNP.

Un aperçu global de l'organisation de la TCSR est offert en (8).

(8) *Aperçu de l'organisation de la TCSR*

TCSR

| | |
|---|---|
| Viol. de contraintes | Régie par |
| ↓ | • le Principe de préservation (limité par le seuil de tolérance) |
| Réparation | • le Principe de minimalité (basé sur la HNP) |
| (insertion / élision) | • la Convention de préséance (basée sur la HNP) |

En somme, tout comme la *Théorie de l'optimalité* (OT; Prince et Smolensky 1993 et McCarthy et Prince 1994), la TCSR est un modèle de contraintes, à la différence que la TCSR est dérivationnelle. Les contraintes interagissent avec les processus et s'appliquent à des niveaux intermédiaires au lieu de ne s'appliquer qu'à des formes de surface comme dans OT.

## 3. LES ADAPTATIONS

Les inventaires segmentaux du peul, la langue la plus répandue géographiquement en Afrique de l'Ouest, sont présentés ci-dessous.

(9) a) inventaire consonantique du peul[3]

| | labial | coronal | | dorsal | laryngal |
|---|---|---|---|---|---|
| | | +antérieur | -antérieur | | |
| occlusives | p / b | t / d | c / ɟ | k / g | |
| implosives | ɓ | ɗ | ʄ | | |
| fricatives | f | s | | | h |
| nasales | m | n | ɲ | ŋ | |
| liquides | | r / l | | | |
| glides | w | | j | | |

b) inventaire vocalique du peul: i, u, ɛ (e), ɔ (o), a

[3] Les alvéo-palatales *c* et *ɟ* sont en fait des affriquées, traditionnellement transcrites comme des occlusives palatales. Nous respectons ici cette convention.

(9b) indique que le peul ne permet pas les voyelles nasales, ce qui est exprimé par le réglage paramétrique négatif en (10a). Malgré cela, les voyelles nasales des emprunts français ne sont presque jamais élidées. On voit en (10b) que, sur 321 cas, il n'y a que 2 cas d'élision. Même l'élision de la nasalité de la voyelle est rare (11% des cas). La nasalité de la voyelle est généralement reconvertie en une consonne nasale, ce qu'on peut observer dans les exemples (10c).

(10) a) *Paramètre*: voyelles nasales? français: oui
(*[-consonantique][+nasal]) peul: **non (contrainte)**

b) Nombre de cas: 326
Nombre d'adaptations: 321 98,5%
$\tilde{v} \rightarrow$ VN: 289 / 321 89%
$\tilde{v} \rightarrow$ V(V): 35 / 321 11%
Nombre d'élisions ($\tilde{v} \rightarrow \emptyset$): 2 0,6%
Nombre de non-adaptations: 3 0,9%

c)

| *français* | | | *peul* |
|---|---|---|---|
| bandit | [b ã di] | → | b an ⁿdi[4] |
| canton | [k ã t õ ] | → | k an t ɔŋ |
| marin | [mar ɛ̃ ] | → | mar ɛŋ |

Les emprunts contiennent aussi souvent des suites de segments mal formés, comme en (11). Par exemple, le mot français *changer* contient une suite de trois segments mal formés pour le peul, *ʃ*, interdit par la contrainte (13), *ã*, interdit par la contrainte (10a), et *ʒ*, interdit par les contraintes (12) et (13).

(11) *$v \rightarrow w$, *$ʒ \rightarrow s$, *$\tilde{v} \rightarrow VN$, *$ʃ \rightarrow s$:

| *français* | | | *peul* |
|---|---|---|---|
| changer | [ ʃã ʒ e] | → | sans -u-dɛ |
| gendarme | [ ʒã darm] | → | san ⁿdarma[5] |
| ventilateur | [ vã tilatœr] | → | wan tilatɔr |

[4] À la strate 1, une consonne nasale prénasalise la consonne suivante si celle-ci est une occlusive voisée (cf. Paradis 1992).

[5] Un noyau nouvellement créé afin de briser une suite consonantique non syllabable est comblé, dans la vaste majorité des cas, par la propagation d'une voyelle ou glide environnante.

(12) *Paramètre*: fricatives voisées? (*[+continu] [+voisé]) français: oui
peul: **non** (*v, *z, *ʒ)

(13) *Paramètre*: fricatives non antérieures? (*[+continu] [-antérieur]) français: oui
peul: **non** (*ʒ, *ʃ)

Pourtant, aucun des segments mal formés n'est élidé. Même quand un segment contrevient à plusieurs contraintes à la fois comme dans le cas de la fricative voisée non antérieure *ʒ, qui viole les contraintes (12) et (13), il n'y a pas élision mais plutôt substitution systématique de *ʒ en *s*, comme on peut le voir en (14a, b). Cette adaptation, où les traits [+voisé] et [-antérieur] sont élidés, est minimale dans l'optique du principe de minimalité ((5)), puisque les stratégies d'élision s'appliquent au niveau le plus bas auquel font référence chacune des deux contraintes (12) et (13), soit le niveau des traits terminaux.

(14) a) *La fricative voisée non antérieure*: *ʒ → s

| *français* | | | *peul* |
|---|---|---|---|
| barrage | [bara ʒ ] | → | bara s |
| collège | [kɔlɛ ʒ ] | → | kɔlɛɛ s |
| journal | [ ʒ urnal] | → | s uurnal |

b) ʒ: Nombre de cas: 63

| | | |
|---|---|---|
| Nombre d'adaptations: | 63 | 100% |
| ʒ → s: | 54 | 85,7% |
| ʒ → z: | 7 | 11,1% |
| ʒ → ʤ: | 2 | 3,2% |
| Nombre d'élisions: 0 | | 0% |
| Nombre de non-adaptations: 0 | | 0% |

En (14b), on peut voir qu'il y a tout de même 7 cas où *ʒ est adapté en *z*, et non en *s*, l'adaptation normale. La fricative voisée *z* commence en fait à être tolérée en peul, ce qui en fait un segment «périphérique» (cf. Paradis et Lebel, ce volume, pour un exposé sur la «périphérie») ou une «importation» dans la terminologie de Haugen (1950). Des exemples sont présentés en (15). L'intérêt de ces exemples, c'est qu'ils nous montrent que les contraintes (12) et (13) sont vraiment deux contraintes séparées, et que *ʒ viole vraiment deux contraintes, puisqu'il y a des cas où la contrainte (12) peut ne pas être satisfaite, c'est-à-dire des cas où le trait [+voisé] est conservé même si la consonne est [+continu].

(15) *Périphérie: dissociation de [-antérieur] mais pas de [+voisé]:*

| *français* | | | *peul* |
|---|---|---|---|
| agence | [a ʒ ãs] | → | a z ãs |
| bougie | [bu ʒ i] | → | buu z i |
| janvier | [ ʒ ãvje] | → | z anvije |

Les malformations syllabiques sont aussi adaptées. Par exemple, en peul, les attaques et les codas branchantes sont interdites, ce qui est exprimé par le réglage paramétrique négatif en (16b).

(16) *Paramètres syllabiques*:

| | | | |
|---|---|---|---|
| a) | i. noyaux complexes? | français: | oui |
| | | peul: | oui |
| | ii. tous les types? | peul: | **non** |
| | -ex.: diphtongues montantes? | peul: | **non** (**wi*, **ji*, etc.) |
| b) | constituants complexes non nucléaires? | français: | oui |
| | | peul: | **non** |

Les mots en (17), (18) et (19) sont non syllabables en peul parce qu'ils contiennent soit une attaque branchante (deux consonnes (CC) en début de mot) soit une coda branchante (CC en fin de mot). Deux choix se présentent alors: élider une des deux consonnes de la suite ou insérer une voyelle dans la suite. Comme on peut le voir, la deuxième option est choisie, tel que prédit par le principe de préservation ((2)). Le segment épenthétique provient du segment vocalique le plus proche, une voyelle ou une glide, et est inséré dans la suite si la sonorité est montante (17) et (18), ou après la suite si la sonorité est descendante (19). Dans tous les cas, la réparation est effectuée au niveau syllabique (insertion d'un noyau), en accord avec le principe de minimalité ((5)), puisque le problème se situe au niveau syllabique.

(17) *Sonorité montante: #CC → CVC*:

| | *français* | | | *peul* |
|---|---|---|---|---|
| a) | place | [ pl as] | → | pal as |
| | tracteur | [ tr aktœr] | → | tar aktɔr |
| | crayon | [ kr ɛjɔ̃] | → | kɛr ɛjɔŋ |
| b) | boisson | [ bw asɔ̃] | → | buw asɔŋ |
| | coiffer | [ kw afe] | → | kuw a:f-a:-dɛ |
| | lieutenant | [ lj øtnã] | → | lij etinaŋ |

(18) *Sonorité montante: CC#→ CVC:*

| *français* | | | *peul* |
|---|---|---|---|
| filtre | [fil tr ] | → | fil tir |
| mètre | [mɛ tr ] | → | mɛɛ tɛr |
| table | [ta bl ] | → | taa bal |
| lettre | [lɛ tr ] | → | lɛɛ tɛr |

(19) *Sonorité descendante: CC#→ CCV:*

| *français* | | | *peul* |
|---|---|---|---|
| carde | [ka rd ] | → | ka rda |
| carte | [ka rt ] | → | ka rt-al [6] |
| force | [fɔ rs ] | → | fɔ rsɔ |

## 4. LES ÉLISIONS

Il existe cependant des cas de malformations syllabiques où il y a perte segmentale, comme dans les exemples (20).

(20) *Cas d'élision: violation du principe du seuil de tolérance ((3))*

| | *français* | | | *peul* | |
|---|---|---|---|---|---|
| a) | voyou | [vw aju] | → | w aju | * wuw aju |
| | voyage | [vw ajaʒ] | → | w aja:s | * wuw aja:s |
| b) | cuivre | [kɥi vr ] | → | ki r i | *ku wir i |
| | pieuvre | [pjœ vr ] | → | piu: r i | *piu: wir i [7] |
| c) | chewing gum | [ʃw iŋgɔm] | → | s iŋᵑgɔm | * siw iŋᵑgɔm |
| d) | cuivre | [kɥ ivr] | → | k iri | * kuw iri |
| | cuisinier | [kɥ izinje] | → | k isiŋᵑgɛ | * kuw isiŋᵑgɛ |

La différence entre les exemples (20), d'une part, et les exemples (17), (18) et (19), d'autre part, c'est que les exemples (20) contiennent tous, en plus d'une malformation syllabique, une malformation segmentale imbriquée dans la malformation syllabique.

---

[6] Le suffixe *-al* est un marqueur de classe nominal. Certains emprunts sont adaptés morphologiquement et prennent alors de tels suffixes.

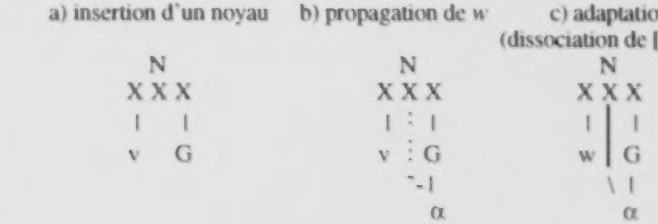
[7] Le *i* final dans *piu:ri* provient d'un marqueur de classe (/piu:r-i/) étant donné que l'article défini est ⁿ*di* (*piu:ri* ⁿ*di* 'la pieuvre').

Par exemple, *voyou* en (20a) contient une suite problématique CC en début de mot, où une des deux consonnes de la suite est une consonne interdite en peul, soit *v*. Même chose pour *cuivre* en (20d), qui commence par une suite CC où la deuxième consonne est la glide palatale arrondie *ɥ*, une glide interdite en peul. Notons que l'élision de cette glide pourrait difficilement être attribuée à un problème de perception phonétique, comme on pourrait le penser à prime abord, puisque, si elle ne cause pas de problème syllabique, comme en (21), elle n'est pas élidée.

| | | | | | |
|---|---|---|---|---|---|
| (21) | (de l') huile | [də lɥ il] | → | di lw il | *dilil |
| | minuit | [mi nɥ i] | → | mi nw i | *mini |

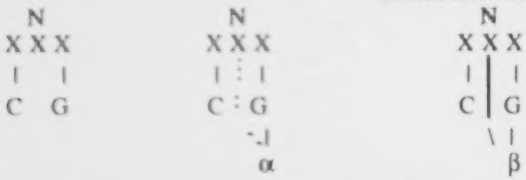
Il n'y a pas de perte de segment en (21), mais plutôt reconversion de *ɥ* en *w*, parce que le problème est simplement segmental, non syllabique. En effet, la glide peut très bien former l'attaque de la deuxième syllabe, et la consonne précédente la coda de la première. Même si ce contact syllabique n'est pas optimal du point de vue de la sonorité, il est permis en peul (Paradis 1992). Par contraste, parce que le problème est à la fois syllabique et segmental en (20), la sauvegarde de tous les segments encadrés serait trop coûteuse. Elle impliquerait plus de deux étapes, ce qui violerait la limite établie par le peul dans le cadre du principe du seuil de tolérance ((3)). Comme on peut le voir en (22) et (23), la préservation des segments mal formés dans les suites non syllabables en (20) nécessiterait 1) l'insertion d'un noyau, comme en (22a) et (23a); 2) la propagation ou l'insertion d'un segment dans le noyau vide, comme en (22b) et (23b), et 3) l'adaptation du segment mal formé lui-même, comme en (22c) et (23c).

(22) **vw* → *wuw*



(23) **Cɥ → Cuw*

a) insertion d'un noyau b) propagation de *ɥ* c) adaptation de *ɥ* (dissociation de Coronal)




C'est pour cette raison que les suites non syllabables en (20) perdent une consonne. Le segment élidé est généralement le segment problématique sur les plans syllabique et segmental. Il peut cependant arriver que ce soit le segment permis qui soit élidé si le segment mal formé constitue, au point de vue de la sonorité, une attaque ou une coda nettement supérieure. C'est le cas de *chewing gum*, où c'est *w* qui est perdu, et non *f*, le segment mal formé, parce les fricatives constituent de meilleures attaques au point de vue de la sonorité que les glides.

On pourrait être tenté de penser en examinant les exemples (20) que la labialité joue un rôle dans les élisions puisque tous les segments élidés dans ces exemples sont des labiales, c'est-à-dire *v*, *w* ou *ɥ*. Cette hypothèse est toutefois infirmée par les exemples (24), où l'on voit que l'élision s'applique aussi à des dentales (*n* ou *d*).

(24) *Élision de dentales*

| *français* | | | *peul* | |
|---|---|---|---|---|
| éponge | [ep[ɔ̃ʒ]] | → | ep[ɔs] | *ep[ɔns]ɔ |
| balance | [bal[ɑ̃s]] | → | bal[as] | *bal[ans]a |
| dimanche | [dim[ɑ̃ʃ]] | → | dim[aːs] | *dim[ans]a |
| essence | [ɛs[ɑ̃s]] | → | ɛs[aːs] | *ɛs[ans]a |
| vacances | [vak[ɑ̃s]] | → | wakk[as] | *wakka[ns]a[8] |
| dépense | [dep[ɑ̃s]] | → | depp[as] | *deppa[ns]a |
| propagande | [prɔpag[ɑ̃d]] | → | pɔrpag[an] | *pɔrpag[and]a |

Dans la plupart de ces exemples, le segment perdu est la consonne nasale, normalement générée suite au décompactage de la voyelle

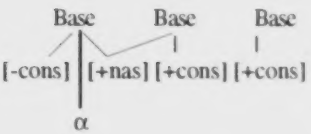
[8] Les occlusives sourdes sont presque toujours géminées à l'intérieur des mots en peul (cf. Paradis 1992: 117 ff.).

nasale comme en (10c). La différence entre les exemples (10c) et (24), c'est que la voyelle nasale dans les exemples (24) est suivie d'une consonne en finale de mot, alors que la voyelle nasale dans les exemples (10c) est elle-même en finale de mot ou suivie d'une consonne syllabable dans une attaque. Décompacter les voyelles nasales n'entraîne donc pas les mêmes conséquences en (10c) qu'en (24). En (24), cela génère automatiquement une suite CC en fin de mot, une suite non syllabable en peul à cause de la contrainte (16b) contre les attaques et les codas complexes. Suite au décompactage, une des deux consonnes, la consonne nasale tirée de la voyelle ou la consonne suivante, doit donc être élidée. Si la consonne finale est une occlusive, comme dans le cas de *propagande*, c'est l'occlusive qui tombe, parce que la nasale forme une coda nettement supérieure à l'occlusive au point de vue sonorité. Si la différence de sonorité est moindre entre les deux consonnes, comme dans le cas d'une nasale et d'une fricative, c'est la consonne problématique au point de vue segmental et syllabique qui tombe, c'est-à-dire la nasale.

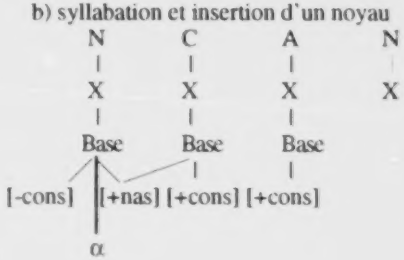
Il y a élision parce que, comme dans les exemples (20), sauvegarder les deux consonnes, impliquerait trop d'étapes par rapport au principe du seuil de tolérance. Il faudrait 1) insérer un noyau comme en (25b); 2) remplir le noyau vide comme en (25c) et 3) dissocier le trait [+nasal] de la voyelle, comme en (25d).

(25) *Dérivation de balance [balãs] → balas (*[balans])*[9]

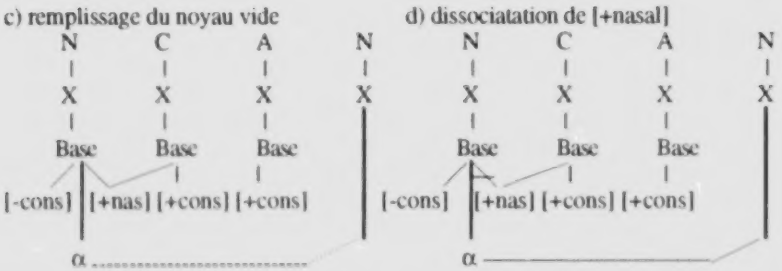
a) forme sous-jacente




b) syllabation et insertion d'un noyau




[9] Il est tenu pour acquis que les voyelles nasales sont constituées de deux segments en forme sous-jacente, c'est-à-dire une voyelle et une consonne partageant un même trait [+nasal] (cf. Paradis et Rose 1995 et, surtout, Paradis et Prunet 1996 pour des arguments).




On peut cependant se demander encore une fois si les élisions en (24), qui visent dans la plupart des cas la nasale, ne résulteraient pas d'un effet phonétique, c'est-à-dire du fait que les suites «voyelles nasales + fricatives» constituent un mauvais contact syllabique (Beddor 1993 et Ohala et Ohala 1993). Cette hypothèse est toutefois rejetée car ces suites sont relativement fréquentes en peul. On les trouve à la fois dans des mots natifs ((26a)) et des emprunts ((26b)).

(26) a) *mots natifs*

| | |
|---|---|
| ni ms -u-dɛ | 'regretter' |
| ja nf -aa-dɛ | 'trahir' |
| le ms -u-dɛ | 'parler beaucoup' |

b) *emprunts*

| *français* | | | *peul* | |
|---|---|---|---|---|
| conseil | [k ɔ̃s ɛj] | → | k ɔns ɛj | *k ɔs ɛj |
| consulat | [k ɔ̃s yla] | → | k ɔns ula | *k ɔs ula |
| changer | [ʃ ɑ̃ʒ e] | → | s ans udɛ | *s as udɛ |
| ingénieur | [ ɛ̃ʒ enjœr] | → | ɛns ɛnjɔr | *ɛs ɛnjɔr |

## 5. CONCLUSION

Nous avons vu que la phonologie joue un rôle certain dans le traitement des emprunts français en peul et que ce traitement est largement prévisible en fonction des principes de la TCSR. Plus précisément, nous avons constaté que 1° la préservation segmentale dans les emprunts est très importante, ce qui est expliqué par le principe de préservation, 2° que les adaptations sont effectuées minimalement en fonction du principe de minimalité et 3° que les rares cas d'élision de segments sont régis par le principe du seuil de

tolérance. Ces résultats sont conformes à ceux que nous avons obtenus dans notre corpus d'emprunts français en kinyarwanda (Paradis et Rose 1995).

Avant de terminer, soulignons un point important pour les neuro et psycholinguistes: les généralisations tirées des corpus d'emprunts que mon équipe et moi avons construits peuvent aussi être tirées de formes suscitées, c'est-à-dire d'emprunts français fictifs que nous avons demandé à nos informateurs peuls d'adapter. Autrement dit, on observe le même comportement dans des formes *off-line* (emprunts réels) que dans les formes *on-line* (emprunts fictifs adaptés par des bilingues). Ceci tend à confirmer que les emprunts (réels) sont non seulement introduits mais aussi «adaptés» par les bilingues (pour des exemples et une discussion, cf. Paradis et LaCharité 1997 et Béland et Paradis 1996).

**REMERCIEMENTS**

Cet article a profité des commentaires de plusieurs collègues au moment de sa présentation orale au colloque «Explorations du lexique», organisé par Julie Auger dans le cadre du congrès de l'*ACFAS*, qui s'est tenu en mai 1996 à l'université McGill. Je remercie particulièrement Denis Dumas, Glyne Piggott, Shana Poplack et Gilian Sankoff. Je suis reconnaissante aux directeurs de publication, Julie Auger et Yvan Rose, pour leurs révisions et suggestions. Nos recherches sont financées par le CRSH #410-94-1296 et le FCAR-équipe #95-ER-2325.

**RÉFÉRENCES**

Beddor, Speeter Patrice. 1993. «The Perception of Nasal Vowels». Dans Marie Huffman et Rena Krakow (réd.). *Nasals, Nasalization and the Velum*. San Diego: Academic Press, pp. 171-196.

Béland, Renée et Carole Paradis. 1996. «Principled Phonological Dissolution in a PPA Case: A comparison between Paraphasias and Loanword Adaptations», Congrès de l'*International Clinical Phonetic and Linguistic Association (ICPLA)*. Munich, septembre.

Haugen, Einar. 1950. «The Analysis of Linguistic Borrowings». *Language* 26:210-231.

McCarthy, John et Alan Prince. 1994. «Prosodic Morphology I: Constraint Interaction and Satisfaction». Dans Geert Booij et J.

van Marle (réd.). *Yearbook of morphology 1993*. Dordrecht: Kluwer, pp. 79-153.

Mougeon, Raymond, Edouard Beniak et Daniel Valois. 1985. *Variation in the Phonological Integration of Loanwords in a Bilingual Speech Community*. Toronto: Center for Franco-Ontarian Studies.

Ohala, John et Manjari Ohala. 1993. «The Phonetics of Nasal Phonology: Theorems and Data». Dans Marie Huffman et Rena Krakow (réd.). *Nasals, Nasalization, and the Velum*. San Diego: Academic Press, pp. 225-249

Paradis, Carole. 1988a. «On Constraints and Repair Strategies». *The Linguistic Review* 6:71-97.

___. 1988b. «Towards a Theory of Constraint Violations». *McGill Working Papers in Linguistics* 5,1:1-43.

___. 1992. *Lexical Phonology and Morphology: The Nominal Classes in Fula*. New York: Garland Publishing.

Paradis, Carole et Darlene LaCharité (réd.). 1993. *Constraint-Based Theories in Multilinear Phonology*: Canadian Journal of Linguistics 38,2.

___. 1996. «Saving and Cost in Loanword Adaptation: Predictions of the TCRS-Phonological Model». *McGill Working Papers in Linguistics* 11:46-84.

___. 1997. «Preservation and Minimality in Loanword Adaptation». *Journal of Linguistics* 33,1. [à paraître]

Paradis, Carole et Jean-François Prunet. 1996. «Nasal Vowels as Two Segments». Université Laval et Université du Québec à Montréal. Ms.

Paradis, Carole et Yvan Rose. 1995. «Préservation et perte segmentale dans les emprunts français en kinyarwanda». Dans Päivi Koskinen (réd.). *Proceedings of the 1995 Annual Conference of the Canadian Linguistic Association*. Toronto: Toronto Working Papers in Linguistics, pp. 411-422.

Poplack, Shana, David Sankoff et Christopher Miller. 1988. «The Social Correlates and Linguistic Processes of Lexical Borrowing and Assimilation». *Linguistics* 26:47-104.

Prince, Alan et Paul Smolensky. 1993. «Optimality Theory: Constraint Interaction in Generative Grammar». Rutgers University et University of Colorado at Boulder. Ms.

Julie Auger et Yvan Rose (réd.). (1997). *Explorations du lexique*. Québec: CIRAL. 229-242.

# Centre et périphérie: les segments anglais en français québécois

*Carole Paradis et Éliane Lebel*
Université Laval

## 1. INTRODUCTION

Les emprunts contiennent souvent des sons inacceptables du point de vue de la langue emprunteuse. Cette langue a alors deux choix: adapter ces sons étrangers ou élargir son inventaire de phonèmes de manière à accepter tels quels ces sons non natifs. Ces sons sont ce qu'on appelle des «importations» ou des non-adaptations (p. ex., le [ɚ] de 'buffer' [bʌfɚ], un segment anglais toléré par de nombreux locuteurs du français québécois (FQ), est une importation). Cependant, les importations ne sont pas toutes «tolérées» au même degré par la langue emprunteuse. Le corpus d'emprunts anglais en FQ construit et analysé par Paradis et Lebel (1994) à partir de données recueillies auprès d'informateurs de la ville de Québec a conduit ces auteures à formuler des prédictions implicationnelles quant aux adaptations et non-adaptations de segments. Par exemple, si les segments [ʧ] ou [ʤ], des segments anglais largement tolérés en FQ, sont adaptés (en l'occurrence, en [ʃ] et [ʒ]) dans un emprunt qui contient aussi les segments anglais [ɚ], [ɹ], [ŋ] et/ou [ɬ], des segments moins tolérés en FQ, il est prédit que ces derniers seront aussi adaptés, c'est-à-dire convertis en un segment natif du français.

L'objectif de notre article est de montrer que les prédictions implicationnelles formulées par Paradis et Lebel (1994) à partir du corpus de la ville de Québec (Qc) sont aussi vérifiées dans le corpus Montréal (Mtl) (composé d'emprunts anglais en FQ de Montréal).[1] Le corpus Mtl présente l'intérêt de contenir un nombre considérablement plus important de non-adaptations que le corpus Qc — même si, comme nous le verrons, ce dernier en contient déjà beaucoup plus que les autres corpus — nous permettant ainsi de vérifier nos prédictions implicationnelles avec plus de certitude. Il est toutefois important de noter que notre article ne constitue pas une comparaison de deux dialectes du FQ dans un but sociolinguistique mais une étude phonologique du traitement des emprunts. Plus généralement, notre étude vise à justifier l'hypothèse d'une

[1] Les emprunts testés auprès des informateurs de Québec et de Montréal proviennent de la même base de données.

«périphérie», le lieu par où entrent les importations, et, enfin, à en préciser l'organisation dans le cerveau de nos informateurs.

Nous tenons pour acquis que 1° les emprunteurs sont des bilingues au sens large du terme (cf. Poplack et al. 1988 pour des arguments sociolinguistiques et Paradis et LaCharité 1996, 1997 pour des arguments phonologiques); cette hypothèse est fortement appuyée par l'existence de non-adaptations; 2° les non-adaptations sont régies par la périphérie, un domaine où les contraintes du «noyau phonologique» sont affaiblies ou éliminées (cf. Chomsky 1986, Itô et Mester 1993/95, Paradis et Lebel 1994).[2]

## 2. DONNÉES ET CADRE THÉORIQUE

Les emprunts des corpus utilisés dans cette recherche ont été recueillis dans divers ouvrages et en discours spontané. Ils ont tous été vérifiés auprès de trois informateurs. La prononciation d'un emprunt par chacun des informateurs a été enregistrée, transcrite en API, informatisée dans une base de données et décortiquée en segments. Ceux-ci ont par la suite été compilés et, finalement, analysés dans le cadre de la *Théorie des contraintes et stratégies de réparation* (TCSR, Paradis 1988a,b).

L'analyse des corpus d'emprunts français en arabe marocain, en kinyarwanda et en peul nous a permis de constater que la préservation segmentale est très forte. Nous pouvons observer en (1) que, sur 8962 malformations, les segments sont préservés (adaptés plutôt qu'élidés) à 90,1%.

[2] Une contrainte «affaiblie» n'est pas une contrainte violée mais une contrainte dont la classe phonologique à laquelle elle s'applique devient tout simplement plus restreinte. Par exemple, une contrainte *x*, au lieu de s'appliquer à toutes les fricatives, ne s'appliquera plus qu'à une fricative donnée (cf. Itô et Mester 1993/95 pour un exposé détaillé sur l'affaiblissement des contraintes). Les contraintes «éliminées» ne sont pas non plus violées mais désactivées, c'est-à-dire que leur domaine d'application n'englobe plus un domaine *y* (cf. Paradis 1988a,b sur la notion de «domaine de contrainte»).

| (1) | Emprunts français en: | | | |
|---|---|---|---|---|
| Nombre de: | *arabe marocain* | *kinyarwanda* | *peul* | Total |
| emprunts | 1547 | 756 | 545 | 2848 |
| formes[3] | 3512 | 2143 | 1036 | 6691 |
| malformations | 3660 | 4444 | 858 | 8962 |
| adaptations | 2981 (81,4%) | 4313 (97,1%) | 785 (91,5%) | 8079 (90,1%) |
| élisions | 230 (6,3%) | 104 (2,3%) | 57 (6,6%) | 391 (4,4%) |
| non-adaptations | 449 (12,3%) | 27 (0,6%) | 16 (1,9%) | 492 (5,5%) |

La TCSR explique cette forte tendance à la préservation des segments par le principe de préservation, en (2).

(2) *Principe de préservation*: L'information segmentale est préservée maximalement, dans les limites du principe du seuil de tolérance.

Comme l'indique le tableau en (1), les élisions sont très rares. Elles sont dues principalement au principe du seuil de tolérance, en (3).

(3) *Principe du seuil de tolérance*:
a) toutes les langues imposent une limite à la préservation segmentale;
b) cette limite est fixée à un nombre d'étapes ou stratégies de réparation *x* à l'intérieur d'un domaine de contrainte.

Un segment emprunté et perçu comme mal formé par la langue emprunteuse est donc adapté plutôt qu'élidé à condition que cette adaptation respecte la limite fixée par la langue quant au seuil de tolérance. Il est aussi posé que cette adaptation doit être minimale.

(4) *Principe de minimalité*: une stratégie de réparation
a) doit s'appliquer au niveau phonologique le plus bas auquel fait référence la contrainte violée
b) en impliquant le moins d'étapes (ou stratégies) possible.

Le niveau phonologique le plus bas auquel fait référence le principe de minimalité est déterminé par la hiérarchie suivante.

---

3 Une «forme» est un emprunt prononcé par un informateur, c'est-à-dire la réalisation concrète d'une entité abstraite, l'emprunt. Cette distinction entre «forme» et «emprunt» est essentielle puisque les informateurs ne connaissent pas tous nécessairement les emprunts qui leur sont présentés ou ne les prononcent pas tous nécessairement de la même manière. Pour plus de détails sur les emprunts et les emprunteurs, cf. Paradis et LaCharité (1996, 1997).

(5) *Hiérarchie des niveaux phonologiques*:
Palier métrique > palier syllabique > squelette > noeud de base > trait non terminal > trait terminal.

## 3. LE CORPUS QUÉBEC

Les inventaires consonantiques et vocaliques de l'anglais et du FQ sont présentés en (6) et (7), respectivement.

(6) Inventaires consonantique et vocalique de l'anglais

| | Labial | Coronal | | | Dorsal | Laryngal |
|---|---|---|---|---|---|---|
| | | [+antér.] | | [-antér.] | | |
| | | [-strid.] | [+strid.] | | | |
| occlusives | p / b | | t / d | | k / g | |
| fricatives | f / v | θ / ð | s / z | ʃ / ʒ | | h |
| affriquées | | | | ʧ / ʤ | | |
| nasales | m | | n | | ŋ | |
| liquides | | | l (ɫ) / ɹ | | | |
| glides | | | | j | w | |

| | | | |
|---|---|---|---|
| i | | | u |
| ɪ | | | ʊ |
| e | ɔ | | o |
| ɛ | | ʌ | ɔ |
| æ | a | | ɑ |

(7) Inventaires consonantique et vocalique du français québécois

| | Labial | Coronal | | Dorsal |
|---|---|---|---|---|
| | | [+antér.] | [-antér.] | |
| occlusives | p / b | t / d | | k / g |
| fricatives | f / v | s / z | ʃ / ʒ | |
| nasales | m | n | ɲ | |
| liquides | | ʁ[4] / l | | |
| glides | | | j / ɥ | w |

| | | |
|---|---|---|
| i | y | u |
| e | ø | o |
| ɛ / ɛ̃ | ɔ / œ / œ̃ | ɔ / ɔ̃ |
| | a | ɑ / ɑ̃ |

La comparaison des inventaires (6) et (7) nous permet de prédire quels seront les segments anglais qui feront l'objet de contraintes en FQ. Ces contraintes, exprimées par les réglages paramétriques négatifs en (8), interdisent les segments anglais entre parenthèses.

(8) Réglages paramétriques de l'anglais et du FQ

| | *anglais* | *FQ* |
|---|---|---|
| Affriquées non antérieures? | oui | non (*ʧ, *ʤ) |
| Interdentales? | oui | non (*θ, *ð) |
| Laryngales? | oui | non (*h) |
| Rétroflexes? | oui | non (*ɹ, *ɚ) |

[4] Bien que *r* soit prononcé comme une uvulaire ou une pharyngale dans de nombreux dialectes du français, cette consonne se comporte phonologiquement comme une coronale à bien des égards (cf. Béland, Paradis et Bois 1993 pour une discussion sur le comportement du *r* en français et en parler aphasique).

| | | |
|---|---|---|
| Vélaires liquides? | oui | non (*ɬ) |
| Vélaires nasales? | oui | non (*ŋ) |

Conformément à ce que l'on a déjà observé dans les corpus en (1), le corpus Qc révèle que la préservation segmentale est importante. On voit dans le tableau en (9) qu'elle correspond à 72% des cas, soit 2521 adaptations sur 3501 malformations.

| (9) Corpus Qc | Total | % |
|---|---|---|
| emprunts | 955 | |
| malformations | 3501 | |
| adaptations | 2521 | 72,0% |
| élisions | 90 | 2,6% |
| non-adaptations | 885 | 25,4% |

Les élisions ne représentent que 2,6% des cas et visent, dans plus de 85% des cas, la laryngale [h], comme dans les exemples (10).

(10) Exemples d'élision de [h] (corpus Qc)

| | | | |
|---|---|---|---|
| *helper* | [hɛɬpɚ] | → | [ɛlpœʀ] |
| *hood* | [hʊd] | → | [ʊd] |
| *skin head* | [skɪnhɛd] | → | [skɪnɛd] |
| *steak house* | [stekhaws] | → | [stekɑws] |

La différence entre le corpus Qc et les autres corpus étudiés réside dans son nombre considérablement plus élevé de non-adaptations. Même si la majorité des malformations y sont adaptées, le corpus Qc compte à lui seul plus de 25% de non-adaptations, comme celles des encadrés en (11).

(11) Exemples de non-adaptations (corpus Qc)

| | | | |
|---|---|---|---|
| *after shave* | [æftɚʃev] | → | [aftɚʃejv] |
| *deal* | [di:ɬ] | → | [di:ɬ] |
| *green* | [gɹi:n] | → | [gɹi:n] |
| *jack* | [ʤæk] | → | [ʤak] |
| *ping pong* | [pɪŋpɑŋ] | → | [pɪŋpɔŋ] |
| *switch* | [swɪʧ] | → | [swɪʧ] |

Ceci n'exclut pas que les segments dits «tolérés» comme ceux en (12) soient ou puissent être adaptés dans certains emprunts.

(12) Exemples d'adaptations de segments tolérés (corpus Qc)

| | | | |
|---|---|---|---|
| *all dressed* | [ɑɜ̈d ɹ ɛst] | → | [ɑld ʁ ɛs] |
| *bill* | [bɪ ɬ ] | → | [bɪ l ] |
| *building* | [bɪldɪ ŋ ] | → | [bɪldɪ ɲ ] |
| *chain saw* | [ ʧ ejnsɑ:] | → | [ ʃ ejnsɑ] |
| *disc jockey* | [dɪsk ʤ ɑki] | → | [ʣɪs ʒ ɔke] |
| *flirt* | [fɬ ɚ t] | → | [fl œʁ t] |

Les segments anglais dits «non tolérés» en FQ, comme ceux des encadrés en (13), sont systématiquement adaptés (sauf pour ce qui est de [h], qui est élidé, comme en (10)).

(13) Exemples d'adaptation de segments non tolérés (corpus Qc)

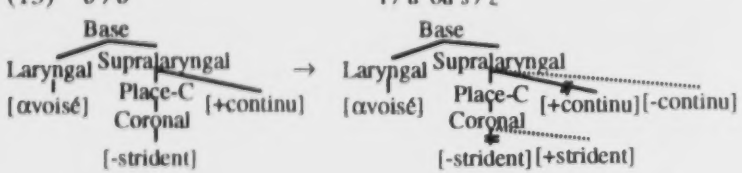
| | | | |
|---|---|---|---|
| *casting* | [k æ stɪŋ] | → | [k a stɪŋ] |
| *make up* | [mek ʌ p] | → | [mek ɔ p] |
| *Thatcher* | [ θ ætʃɚ] | → | [ t atʃɚ] |
| *that's it* | [ ð ætsɪt] | → | [ d atsɪt] |

Nous avons vu en (8) que les adaptations sont dues à des réglages paramétriques négatifs. Nous voyons en (14) quelle est précisément la combinaison de traits prohibée dans le cas des interdentales.

(14) Consonne [-strident] [+continu]? anglais: oui
FQ: non (*θ, *ð)

Le principe de minimalité prédit qu'un des deux traits de la combinaison problématique, soit [+continu] ou [-strident] devra être élidé, comme illustré en (15).

(15) *θ / ð* *t / d* ou *s / z*



Comme en espagnol mexicain et dans beaucoup d'autres langues, c'est le trait [+continu] qui est élidé en FQ, ce qui produit [t] ou [d], selon le cas. Pourquoi cette préférence alors que, en français européen, on préfère élider le trait [-strident] et produire ainsi [s] ou [z]? À cause d'une convention sociale (cf. Paradis et LaCharité 1997). Toutefois, dans les deux cas, l'adaptation est minimale puisqu'elle s'applique à un des deux traits de la combinaison prohibée. En résumé, les segments anglais dans le corpus Qc sont maximalement préservés, c'est-à-dire adaptés plutôt qu'élidés, et adaptés minimalement lorsqu'adaptés, un effet des principes (2) et (4).

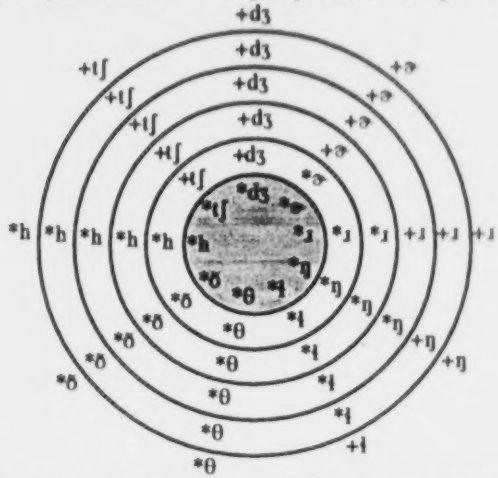
Cependant, comment traiter les non-adaptations, dont le nombre élevé en fait le principal intérêt du corpus Qc par rapport aux autres corpus en (1)? Clairement, ces segments sont différents des segments systématiquement adaptés comme les interdentales [θ] et [ð], mais ils sont aussi différents des segments «natifs» du FQ ([p], [b], [t], etc.). Par ailleurs, comment rendre compte du fait que les segments tolérés ne le sont pas tous au même degré? Par exemple, les statistiques en (16) indiquent que [ʤ] et [ʧ] sont presque toujours tolérés (respectivement 92,4% et 91,7% des cas), mais que [ŋ] et [ɬ] le sont nettement moins (respectivement 14,7% et 1,6% des cas).

(16) % des non-adaptations par segment (corpus Qc)

| | *total des cas* | *total des non-adaptations* | |
|---|---|---|---|
| [ʤ]: | 144 | 133 | 92,4% |
| [ʧ]: | 180 | 165 | 91,7% |
| [ɹ]: | 1027 | 311 | 30,1% |
| [ŋ]: | 190 | 28 | 14,7% |
| [ɬ]: | 448 | 7 | 1,6% |

Un modèle périphérique gradué, comme celui proposé par Itô et Mester (1993/95) et repris par Paradis et Lebel (1994) permet de rendre compte des faits. Le lexique y est représenté par un noyau, la partie en gris en (17), où sont activées toutes les contraintes de la langue. Dans cette partie du lexique, tous les segments anglais sont interdits. Le noyau est entouré de cercles concentriques qui représentent la périphérie. Certains segments de l'anglais y sont tolérés en FQ, ce qui est indiqué par des «plus» en (17). Par exemple, dans le premier cercle, on voit qu'il n'y a que les affriquées [ʧ] et [ʤ] qui sont tolérées. Dans les cercles suivants, on remarque que [ɹ] fait son apparition, suivi de [ɹ], [ŋ] et, finalement, de [ɬ]. Les interdentales [θ] et [ð] et la laryngale [h] ne sont jamais permises, ni dans le noyau ni dans la périphérie des informateurs du corpus Qc.

(17) Représentation de la périphérie du corpus Qc



Ce modèle a l'avantage de faire des prédictions implicationnelles quant à la cooccurrence de segments étrangers dans un emprunt. Les prédictions pour le corpus Qc sont présentées en (18).

(18) Prédictions pour le corpus Qc: à l'intérieur d'un même emprunt
a) si [ʧ] et [ʤ] sont adaptés, [ɚ], [ɹ], [ŋ] et [ɬ] sont adaptés.
b) si [ɚ] est adapté, [ɹ], [ŋ] et [ɬ] sont adaptés.
c) si [ɹ] est adapté, [ŋ] et [ɬ] sont adaptés.
d) si [ŋ] est adapté, [ɬ] est adapté.
e) si [ɬ] est non adapté, [ʧ], [ʤ], [ɚ], [ɹ] et [ŋ] sont non adaptés.

## 4. LE CORPUS MONTRÉAL

Nous avons voulu vérifier si les prédictions implicationnelles dégagées du corpus Qc allaient être les mêmes pour un corpus comme le corpus Mtl, qui devait logiquement contenir un nombre encore plus important de non-adaptations, soit 44,3%, comme indiqué en (19), après compilation des statistiques.

| (19) Corpus Mtl | Total | % |
|---|---|---|
| emprunts | 955 | |
| malformations | 3228 | |
| adaptations | 1721 | 53,3% |
| élisions | 78 | 2,4% |
| non-adaptations | 1429 | 44,3% |

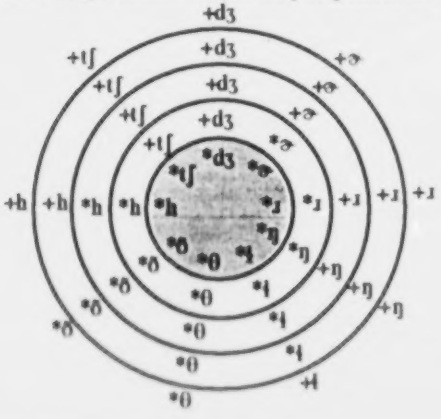
Les non-adaptations dans le corpus Mtl sont plus de 20% plus nombreuses que dans le corpus Qc, qui contient déjà beaucoup plus de non-adaptations que les autres corpus en (1) et, cela va sans dire, que ce que l'on a observé dans une étude préliminaire des emprunts anglais en français européen (Tardif 1996), où le taux de non-adaptation est très faible.[5] Le corpus Mtl permet donc de vérifier avec beaucoup plus de certitude 1° la possibilité de faire des prédictions implicationnelles et 2° la nature de ces prédictions dans les corpus anglais-français. En comparant les chiffres obtenus dans le corpus Mtl avec ceux obtenus dans le corpus Qc (entre parenthèses en (20)), on constate que les segments, [ɹ], [ŋ] et [ɬ], en particulier, sont beaucoup plus tolérés à Montréal qu'à Québec.

(20) % des non-adaptations par segment (corpus Mtl)

| | *total des cas* | *total des non-adaptations* | |
|---|---|---|---|
| [ʤ]: | 132 | 128 | 97,0% (Qc: 133/144 92,4%) |
| [ʧ]: | 173 | 169 | 97,7% (Qc: 165/180 91,7%) |
| [ɹ]: | 940 | 593 | 63,1% (Qc: 311/1027 30,1%) |
| [ŋ]: | 174 | 84 | 48,3% (Qc: 28/190 14,7%) |
| [h]: | 68 | 10 | 12,8% (Qc: 0/77 0%) |
| [ɬ]: | 398 | 33 | 8,3% (Qc: 7/448 1,6%) |

La périphérie tirée du corpus Mtl, et représentée en (21), comporte moins de domaines (cercles) que celle dégagée du corpus Qc.

(21) Représentation de la périphérie du corpus Mtl




[5] Par exemple, [ɹ], qui est un segment relativement toléré en FQ, est adapté dans 100% des 109 occurrences du corpus français européen.

BEST COPY AVAILABLE

On peut également noter l'apparition d'un nouveau segment, la laryngale [h], qui n'est pas du tout tolérée dans le corpus Qc, mais qui commence à l'être dans le corpus Mtl. Cette organisation nous permet de faire les prédictions en (22).

(22) Prédictions pour le corpus Mtl: à l'intérieur d'un même emprunt
a) si [ʧ] et [ʤ] sont adaptés, [ɚ], [ɹ], [ŋ], [h] et [ɬ] sont adaptés;
b) si [ɚ], [ɹ] ou [ŋ] sont adaptés, [ɚ], [ɹ], [ŋ], [h] et [ɬ] sont adaptés;
c) si [h] est élidé, [ɬ] est adapté;
d) si [ɬ] est non adapté, tous les autres segments sauf [θ] et [ð] sont non adaptés.

Si nous comparons (18) et (22), nous pouvons penser que les prédictions pour le corpus Mtl sont un peu différentes de celles dégagées du corpus Qc, surtout pour ce qui est des domaines les plus périphériques. Cependant, aucune des prédictions ne se contredit. Les prédictions en (18b-d) sont simplement fusionnées dans le corpus Mtl parce que les segments [ɚ], [ɹ] et [ŋ] du corpus Mtl sont tolérés au même degré. Autrement dit, la différence entre les deux corpus réside seulement dans une plus longue gradation pour le corpus Qc. À quelques rares exceptions près, les prédictions en (18) et (22) sont confirmées par les données de nos corpus (cf. Paradis et Lebel 1994 pour une discussion de deux contre-exemples).

## 5. DISCUSSION SUR LA PÉRIPHÉRIE

En somme, l'hypothèse d'une «périphérie» graduée est essentielle pour rendre compte des non-adaptations et permet, en outre, de faire des prédictions implicationnelles importantes. Toutefois, on peut se demander ce qu'est un segment «périphérique». Autrement dit, qu'est-ce qui le distingue des autres segments et, finalement, comment peut-il être appris? Paradis et LaCharité (1996, 1997) ont proposé des critères à cet effet, que nous répétons ici. Un segment périphérique:

1) est limité surtout à la catégorie nominale (les emprunts étant principalement des noms). Il n'est jamais contenu dans un mot fonctionnel (préposition, etc.) parce que ce genre de mots ne fait pas l'objet d'emprunt (cf. Vinet 1996:166);
2) ne sert pas à former des mots nouveaux;
3) est moins fréquent lexicalement que la majorité des autres segments;

4) alterne souvent avec un autre segment considéré comme moins marqué du point de vue du noyau central de la langue;
5) est souvent limité à des mots ou à des positions syllabiques spécifiques. Par exemple, le [ʒ] n'apparaît pas à l'initiale de mot pour la majorité des locuteurs anglophones;
6) est souvent contenu dans un mot qui ne subit pas (tous) les processus phonologiques réguliers. Par exemple, [ʧ] dans 'beach' [biːʧ] est contenu dans un mot qui ne subit pas la règle obligatoire de relâchement des voyelles hautes en syllabe fermée (*[bɪːʧ]) en FQ;
7) apparaît souvent en cooccurrence avec d'autres segments qui possèdent les caractéristiques qui viennent d'être énumérées.

Il est donc possible de reconnaître un segment périphérique. Cependant, la notion de périphérie n'équivaut-elle pas au fond à poser une phonologie des emprunts? Nous rejetons l'hypothèse de deux phonologies distinctes, une native et une des emprunts, pour de nombreuses raisons. Premièrement, la plupart des emprunts finissent par être complètement adaptés, c'est-à-dire à faire partie de la phonologie centrale de la langue emprunteuse. Ces emprunts, une fois intégrés au lexique central, cessent d'être perçus comme des mots anglais par l'emprunteur. Ils subissent les processus phonologiques et morphologiques réguliers de la langue emprunteuse. Selon l'hypothèse de deux phonologies, un même emprunt qui contiendrait des segments anglais adaptés et non adaptés devrait se retrouver à deux endroits, la phonologie native et celle des emprunts. Comme le montre Paradis (1995), il va sans dire qu'un tel traitement serait extrêmement redondant. Par exemple, il faudrait poser une règle d'affrication des occlusives dentales dans la phonologie native et une autre dans la phonologie des emprunts (ex.: le mot anglais *disc jokey* [d̲ɪskʤɔkɪ], qui est réalisé [ʤɪsʒɔke] en FQ) et ainsi de suite pour tous les processus.

Deuxièmement, la périphérie ne contient pas de contraintes nouvelles. Si nous acceptons l'idée qu'une contrainte n'est que le refus d'une certaine complexité, nous pouvons dire que la périphérie est tout simplement plus permissive, c'est-à-dire qu'elle est régie par moins de contraintes. Par exemple, on ne s'attendrait pas à ce que la périphérie d'une langue refuse les attaques branchantes, une structure complexe, si la phonologie centrale de cette même langue les accepte. Cette généralisation échappe complètement à l'hypothèse de deux phonologies parallèles. Troisièmement, si nous acceptons cette prémisse, à savoir que la périphérie n'est que l'affaiblissement ou

l'élimination des contraintes de la phonologie centrale, nous pouvons faire des prédictions implicationnelles quant aux non-adaptations, que ne permettrait pas de faire l'hypothèse de deux phonologies distinctes. Quatrièmement, les segments périphériques (encadrés) ne sont pas limités aux emprunts anglais ((23)), ni aux emprunts tout court. En effet, on peut les retrouver dans des interjections ((24)) et des onomatopées ((25)), entre autres. En (25), on note que, pour beaucoup de locuteurs montréalais, l'affrication, un processus normalement obligatoire, ne s'appliquent pas, ce qui confirme le caractère marginal (périphérique) des onomatopées.[6]

(23) Exemples d'emprunts à d'autres langues que l'anglais contenant des segments périphériques

| | | | |
|---|---|---|---|
| *djellaba* | (arabe) | → | [ʤɛlaba] |
| *djinn* | (arabe) | → | [ʤɪn] |
| *dolce* | (italien) | → | [dɔlʧe] |
| *pichounette* | (provençal) | → | [piʧunɛt] |
| *tchador* | (arabe) | → | [ʧadɔʀ] |

(24) Interjections avec segments périphériques

| | | | |
|---|---|---|---|
| *atchoum!* | (FQ) | → | [aʧœm] |
| *ciao!* | (italien) | → | [ʧaw] |
| *tchin tchin!* | (cantonais) | → | [ʧɪnʧɪn] |
| *tchou tchou!* | (anglais) | → | [ʧu:ʧu:] |

(25) Onomatopées: ils peuvent échapper à des processus obligatoires

| | | | |
|---|---|---|---|
| *beding bedang* | → | [bədɪŋbədaŋ] | (*[bədʒɪŋbədaŋ]) |
| *ding dong* | → | [dɪŋdɔŋ] | (*[ʤɪŋdɔŋ]) |

[6] Cela ne veut pas dire qu'aucun Montréalais n'affrique dans les éléments périphériques. Cela veut dire qu'il est possible pour un grand nombre de locuteurs de ne pas appliquer dans un grand nombre d'éléments périphériques, sinon dans tous, un processus autrement obligatoire au niveau lexical. Ceci suffit en soi à faire ressortir le caractère spécial de la périphérie. En ce qui a trait à l'affrication, cela est particulièrement vrai des Montréalais, un peu moins des locuteurs de la ville de Québec, qui tendent à affriquer davantage dans tous les domaines — lexical, postlexical et périphérique.

| | | | |
|---|---|---|---|
| *ting* | → | [t ɪŋ] | (*[ts ɪŋ]) |
| *tût tût* | → | [t yt t yt] | (*[ts yt ts yt]) |

Un nombre important de suites consonantiques que l'on trouve dans les mots savants doivent aussi être considérées comme «périphériques» en ce que ces suites sont limitées à cette catégorie de mots et sont perçues comme mal formées ou marginales par la très vaste majorité des locuteurs. Par exemple, la suite [gn] est limitée à trois mots savants (*gnome*, *gnomon* et *gnose*) et leurs dérivés. Lorsque l'on demande aux locuteurs québécois s'ils pensent que *gnose* est français ou provient d'une langue étrangère, la majorité choisissent la deuxième option. Que dire maintenant des suites [ft] dans *phtisie* et [rkt] dans *infarctus*?

Il s'ensuit que la périphérie ne peut pas être considérée comme l'équivalent d'une phonologie des emprunts.

## 6. Conclusion

En somme, cet article a montré que la notion de «périphérie» est essentielle pour rendre compte des «importations» (non-adaptations). Il a aussi montré qu'une périphérie graduée permet de faire des prédictions implicationnelles importantes par rapport aux importations en FQ et que la nature de ces implications est la même dans les deux corpus FQ. Enfin, il a été montré qu'un segment périphérique peut être distingué d'un segment non périphérique sur la base de critères bien définis. Nous soutenons aussi que la «périphérie» n'équivaut pas à une phonologie des emprunts, une hypothèse que nous rejetons pour de nombreuses raisons, et que la périphérie est indépendamment justifiée.

## Remerciements

Notre article a profité des commentaires de plusieurs collègues lors de sa présentation orale au colloque «Explorations du lexique», organisé par Julie Auger dans le cadre du congrès de l'*ACFAS*, qui s'est tenu en mai 1996 à l'université McGill. Nous remercions particulièrement Denis Dumas, Glyne Piggott, Shana Poplack et Gilian Sankoff. Nous sommes aussi reconnaissantes à Julie Auger et Yvan Rose pour leurs révisions et suggestions. Nos recherches sont financées par le CRSH #410-94-1296 et le FCAR #95-ER-2325.

## RÉFÉRENCES

Béland, Renée, Carole Paradis et Monique Bois. 1993. «Constraints and Repairs in Aphasic Speech: A Group Study». Dans Carole Paradis et Darlene LaCharité (réd.). *Constraint-Based Theories in Multilinear Phonology.* Canadian Journal of Linguistics 38,2:279-302.

Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin and Use.* New York: Praeger.

Itô, Junko et R. Armin Mester. 1993/1995. «Japanese Phonology: Constraint Domains and Structure Preservation». Dans John Goldsmith (réd.). *The Handbook of Phonological Theory.* Cambridge, MA: Blackwell, pp. 817-838.

Paradis, Carole. 1988a. «On Constraints and Repair Strategies». *The Linguistic Review* 6:71-97.

___. 1988b. «Towards a Theory of Constraint Violations». *McGill Working Papers in Linguistics* 5,1:1-43.

___. 1995. «Native and Loanword Phonology as One: The Role of Constraints». Dans Kjell Elenius et Peter Branderud (réd.). *Proceedings of the XIII$^{th}$ International Congress of Phonetic Sciences.* Stockholm: Royal Institute of Technology and Stockholm University, pp. 74-81.

Paradis, Carole et Darlene LaCharité. 1996. «Saving and Cost in Loanword Adaptation: Predictions of the TCRS-Phonological Model». *McGill Working Papers in Linguistics* 11:46-84.

___. 1997. «Preservation and Minimality in Loanword Adaptation». *Journal of Linguistics* 33,1. [à paraître]

Paradis, Carole et Caroline Lebel. 1994. «Contrasts from Segmental Parameter Settings in Loanwords: Core and Periphery in Quebec French». Dans Carrie Dyck (dir.). *Proceedings of the MOT Conference on Contrast in Phonology* 13. Toronto: Toronto Working Papers in Linguistics, pp. 75-95.

Poplack, Shana, David Sankoff et Christopher Miller. 1988. «The Social Correlates and Linguistic Processes of Lexical Borrowing and Assimilation». *Linguistics* 26:47-104.

Tardif, Benoît. 1996. «Recherche sur l'adaptation des emprunts anglais en France». Travail de session, cours LNG 64495, Université Laval.

Vinet, Marie-Thérèse. 1996. «Lexique, emprunts et invariants: une analyse théorique des anglicismes en français du Québec». *Revue québécoise de linguistique* 24:165-181.

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 243-251.

# La prononciation des mots anglais en québécois: le problème particulier des anglophones

*Marc Picard*
Université Concordia

## 1. Introduction

Une des tâches majeures dans l'apprentissage d'une langue seconde consiste évidemment à essayer de se conformer le plus possible à la prononciation des locuteurs autochtones. Puisque plusieurs des difficultés phonologiques auxquelles les apprenants sont confrontés à cet égard sont de nature générale, voire universelle, comme le sont d'ailleurs les stratégies utilisées pour les surmonter, elles sont assez bien connues du fait qu'elles aient fait l'objet d'un bon nombre d'études.[1]

En revanche, il existe un problème plus particulier qui semble avoir échappé à l'attention des chercheurs dans ce domaine. Il s'agit de l'acquisition d'une langue seconde qui contient un nombre important de formes (incluant des noms propres de toutes sortes) qui proviennent de la langue première de l'apprenant, et qu'il doit s'astreindre à 'dénativiser' ou 'dénaturaliser', pour ainsi dire. J'aimerais donc examiner le problème qu'ont les anglophones à réprimer leur tendance naturelle à prononcer à l'anglaise les lexèmes québécois qui proviennent de leur langue maternelle, ce qui nous amènera, par le fait même, à jeter un coup d'oeil sur la façon dont les Québécois eux-mêmes adaptent les emprunts de l'anglais.

## 2. Le problème des noms propres anglais

En tant que francophones, et encore plus en tant que linguistes francophones, nous sommes tous conscients du nombre élevé de mots anglais d'emprunt relativement récent qui font partie de notre vocabulaire quotidien: *hot dog*, *smoked meat*, *hamburger*, *relish*, *ketchup*, *ratchet*, *brake/break*, *steering*, *windshield*, *hood*, *dash*, *tire*,

[1] Pour obtenir une vue d'ensemble des différentes analyses qui ont été proposées à cet effet, voir, par exemple, Picard & Nicol (1982) et Paradis & LaCharité (1997).

*jack*, *plug*, *fucké*, *flyé*, *speedé*, *lucké* (et son dérivé de notre cru *badlucké*), *loose*, *slack*, *rough*, *tough*, *track*, *zipper*, *lighter*, *cartoon (de cigarettes)*, *van*, *job*, *chum*, etc. Je crois que ceci tient en grande partie du fait que nos parents et éducateurs se sont tellement acharnés à nous les proscrire.

Mais à côté de ces quelques centaines de soi-disant 'anglicismes' qui nous sont si familiers, il y a un autre phénomène dont les Québécois, et même les linguistes québécois, paraissent avoir beaucoup moins conscience. Il s'agit littéralement des milliers de noms propres anglais que les francophones nord-américains sont constamment appelés à prononcer. Par exemple, il y a des noms de produits de tous genres (voitures, nettoyants, céréales, appareils ménagers, articles de toilette, etc.), des noms géographiques (états, provinces, régions, villes, quartiers, rues, etc.), des noms d'équipes sportives professionnelles et amateures (hockey, baseball, football, basketball, etc.), etc. Il y a aussi le nombre faramineux de noms de personnes dans les domaines de la politique, du sport, de la vie artistique, de la science, de la littérature, etc. qui doivent être francisés continuellement dans la vie de tous les jours, les *Elvis Presley*, *Madonna*, *Beatles*, *Gordie Howe*, *Eric Lindros*, *Michael Jackson*, *Preston Manning*, *Bill Clinton*, *John Major*, *Hemingway*, *Shakespeare*, *Chomsky*, etc.

À ce propos, j'aimerais citer un extrait d'une lettre parue il y a quelques années dans un hebdomadaire (ou était-ce un mensuel?) qui s'appelait *l'aut'journal* (sic). En fait, je crois que c'est après avoir lu cette lettre que j'ai commencé à prendre conscience de l'ampleur du phénomène des noms propres anglais en québécois. Ça s'intitule *French Québec?*, et en voici deux passages:

> Au lever, *White Swan*, *Close-Up* et *Old Spice* s'occupent de ma toilette de corps. J'hésite ensuite, à la table, entre les *Corn* et les *Frosted Flakes*. Je laisse *Cat Chow* à Minou, et Marie me prie de rapporter quelques *Light Days* à mon retour. J'en profite pour me procurer la poudre *Arctic Power* au passage: elle nettoiera mes vêtements pendant que je m'offrirai une *Golden* ou la nouvelle et très québécoise *Labatt Dry* [. . .]
> Je ne dois pas oublier l'anniversaire d'Olivier. Je passe donc chez *Toys-R-Us*. Une petite fringale me prend. Opterai-je pour *Burger King* ou *McDonald's*? N'étant pas totalement masochiste, je bifurque vers *Red Lobster*, me doutant bien

> toutefois que je mords au même hameçon. Puis c'est la sucrerie au *Dairy Queen*, sinon au *Dunkin' Donuts*. Enfin, je ramasse le papier teint commandé au *Colour your World* et le costume acheté au *Big Steel*. Ah oui, les *Good Year* chez *Firestone* aussi.

Évidemment, il y a lieu de se demander si on peut considérer toutes ces formes au même niveau que les emprunts lexicaux. Je crois que la réponse est 'oui' dans la mesure où il faut les faire passer par le même processus que tous les autres lexèmes anglais pour les naturaliser, c'est-à-dire pour faire en sorte qu'ils se conforment aux règles et contraintes phonologiques du québécois. Autrement dit, ce sont les mêmes stratégies adaptatives, telles la *substitution* (A /θ ð/ > Q /t d/), la *suppression* (A /h/ > Q Ø), et l'*adoption* (A /ɹ/ > Q /ɹ/ pour plusieurs locuteurs), qui sont employées dans les deux cas d'absorption de segments étrangers, à savoir les lexèmes et les noms propres.

Il est donc clair que tous les francophones de l'Amérique du Nord sont sans cesse appelés à prononcer des mots anglais, et pas toujours les mêmes, en plus, car il y a un afflux constant d'anglophones qui se présentent sur la scène politique, artistique, sportive, juridique ou scientifique, et dont on se voit contraints de prononcer les noms à coeur de jour. Bref, pour n'importe quel locuteur autochtone du français d'ici, *y a rien là*!

Cependant, pour les anglophones, ceci peut présenter un véritable casse-tête car ils doivent non seulement réprimer un réflexe naturel de prononcer des mots qui leur sont très familiers de la façon dont ils les ont toujours prononcés dans leur langue maternelle, mais ils doivent aussi apprendre un nombre impressionnant de règles d'adaptation et d'exceptions, surtout en ce qui a trait aux noms propres. Sans chercher à en dresser une liste exhaustive, j'aimerais tout de même tenter de donner une idée de la complexité de la tâche.

### 3. La prononciation: règles et exceptions

Commençons par un fait vécu. Il y a quelque temps, une femme dans une auto m'interpelle au centre-ville de Montréal, et me demande où se trouve la rue /mfɬtən/. N'ayant pas détecté d'accent anglais dans son parler, je l'aurais probablement prise pour une francophone si elle n'avait pas prononcé *Milton* de cette façon, c'est-à-dire avec un /l/ vélarisé et l'accent tonique sur la première syllabe.

Ce sont là deux modifications fondamentales que n'importe quel francophone effectuerait spontanément, me semble-t-il, de sorte qu'on aurait de fortes chances d'entendre ce mot prononcé /mɪltə̃n/ — et même /miltə̃n/ pour ceux dont la règle de relâchement des voyelles fermées se limite aux syllabes accentuées (cf. Picard 1979, 1983, 1996) — dans un contexte français.

Si on se penche maintenant sur la qualité de la voyelle de la dernière syllabe d'une forme comme *Milton* — et il y en a une flopée d'autres du même genre comme *bacon*, *muffin*, *Edmonton*, *Houston*, *Manhattan*, *Moncton*, *Martin*, *Simpson*, etc. — on se rend compte que celle-ci demeure effectivement inchangée.[2] Et pourquoi pas, car il n'existe plus de contrainte contre la présence de schwa en syllabe accentuée en québécois.

Remarquez que ceci n'a pas toujours été le cas en français. Ainsi, Schane (1968) fait remarquer que "[u]nlike the other vowels, /ə/ never is found in tonic position" (19), et à ceux qui offriraient comme contre-exemple une forme comme *dites-le*, il répond que "this particular example is not an instance of stress within the word . . . but is due to the position of *le* within the phrase" (135). En réalité, cependant, il semble qu'il n'y ait plus guère que les dictionnaires qui fassent une distinction entre la voyelle centrale mi-ouverte non arrondie /ə/ de *brebis* et la voyelle antérieure mi-fermée arrondie /œ/ de *neuf*. Si on se fie aux descriptions de Grammont (1914), Fouché (1959), Dell (1973) ou Tranel (1987), par exemple, on voit qu'en français hexagonal les deux voyelles se prononcent /œ/. En Amérique, par contre, c'est clairement le schwa qui a absorbé l'autre, du moins lorsque cette dernière est brève, de telle sorte qu'on entendra communément /bʀəbi/ et /nəf/.

À première vue, on pourrait croire que le maintien intégral de la

[2] Il est possible que certains locuteurs lui substituent la voyelle antérieure arrondie relâchée /ʊ/ mais ceci reste à confirmer. Bien que tous les locuteurs autochtones que j'ai consultés se défendent bien d'avoir une telle voyelle dans ce genre d'emprunt, la plupart d'entre eux ne semblent pas trouver une prononciation comme /mɪltʊ́n/ tout à fait insolite comme le serait /mɪltɔ́/, par exemple. Par contre, la substitution d'une voyelle postérieure mi-ouverte au schwa anglais, qui semble prévaloir en Europe, ne me semble pas très commune ici mais encore là faudrait-il examiner la situation de plus près.

dernière syllabe des mots comme *Milton* serait un avantage pour les anglophones mais ceci n'est pas le cas puisque l'anglais standard nord-américain ne tolère pas le schwa en syllabe accentuée sauf devant /ɹ/. En somme, le dilemme se présente comme suit: si l'accent tonique est transposé à la dernière syllabe, ce qui, à mon avis, est une condition sine qua non de l'intégration pleine et entière d'un mot anglais en français, la voyelle devient alors difficile à prononcer. Dans les circonstances, la meilleure solution de rechange serait probablement de faire comme les Français, et de substituer au schwa une voyelle postérieure mi-fermée de telle sorte que Milton et les autres riment avec *tonne*, *bonne*, *sonne*, etc. Enfin, quelle que soit la qualité de la voyelle dans ce contexte, un anglophone doit savoir que ça en prend absolument une, ce qui signifie que des formes telles que *Eaton* /i:tn/ et *Garden* /gaɹdn̩/ (comme dans *Madison Square Garden* et *Maple Leaf Gardens*) qui se terminent par un /n/ syllabique doivent obligatoirement subir une épenthèse, et se prononcer /itən/ et /gaɹdən/ (ou encore /gaʀdən/) .

Une forme comme *Milton* nous démontre encore autre chose, à savoir qu'une séquence VN$ demeure telle quelle en français moderne, c'est-à-dire que l'époque où elle devait se transformer en Ṽ$ semble tout à fait révolue. Ainsi, des emprunts comme *van*, *can*, *drum*, *fun*, *chum*, *candy*, *sundae*, *pantry* se prononcent comme si on les écrivait *vanne*, *canne*, *dromme*, *tchomme*, *cannedé*, *sonnedé*, *pannetré*. Le problème, cependant, est qu'il y a un bon nombre d'exceptions parmi les noms propres. Prenons, par exemple, les noms de ville *Boston* et *Austin* qui, phonétiquement parlant, ont tout sauf le /b/ en commun. Bien qu'*Austin* soit francisée à /ɑ:stən/ (sauf lorsqu'il s'agit de l'auto anglaise qu'on connaissait sous le nom d'/ʌstẽ/), *Boston* ne se prononce jamais /bɑ:stən/ mais toujours /bʌstɔ̃/ en français québécois.

De même, on pourrait s'attendre à ce que le /væn/ initial de *Vanderbilt* se retrouve également dans *Vancouver* mais cette dernière forme se voit immanquablement francisée à /vãkuvɛ:ʀ/. Ou encore, étant donné le /sæn/ de *San Diego* et de *San Jose*, on aimerait pouvoir prédire cette même séquence phonologique dans la tout aussi californienne *San Francisco*. Mais non, on prononce toujours cette dernière /sãfʀãsisco/. Parmi les autres noms de villes canadiennes et américaines qui contiennent exceptionnellement des voyelles nasales,

on retrouve *Toronto*, *Cincinnati*, *Los Angeles*, *Atlanta* et *Saint-Louis*. Enfin, il ne faut surtout pas oublier de dire /lɔ̃:d(ʀ)/ si on parle de la *London* anglaise, et /lʌndən/ si on fait référence à sa jumelle canadienne.

Une autre question qu'il faut se poser si on est anglophone est la suivante: quand est-ce que j'adapte, et quand est-ce que je traduis? Par exemple, *New Brunswick*, *Nova Scotia*, *New Orleans*, *New England* et *New Mexico* deviennent obligatoirement *Nouveau-Brunswick*, *Nouvelle-Écosse*, *Nouvelle-Orléans*, *Nouvelle-Angleterre et Nouveau-Mexique* mais on n'entendra jamais parler de *Nouveau-York*, du *Nouveau-Jersey* ou du *Nouveau-Hampshire*. Il en va de même pour les noms que portent les très nombreuses équipes sportives professionnelles. Si on n'entend jamais parler des *Feuilles d'Érable* de Toronto, des *Faucons Noirs* de Chicago, des *Bas Blancs* de Chicago, des *Ailes Rouges* de Détroit ou des *Bombardiers Bleus* de Winnipeg, comment se fait-il que les *Pittsburgh Penguins*, les *Pittsburgh Pirates*, les *Atlanta Braves*, les *BC Lions* et le *Colorado Avalanche* deviennent les /pẽgwẽ/, les /piʀæt/, les /bʀæv/, les /lyɔ̃/ et l'/ævælãš/? Peut-être parce que les mots se ressemblent tellement dans les deux langues? Alors, comment expliquer que les *Saints* de la Nouvelle-Orléans et les *Giants* de San Francisco ne deviennent jamais les /sẽ/ ni les /žeã/? C'est à n'y rien comprendre!

Voici une autre question que les anglophones doivent certainement se poser: qu'est-ce que les Québécois font exactement avec les /ɹ/ de l'anglais? S'ils se transformaient tous à /ɾ/ ou à /ʀ/, il s'agirait uniquement de se consacrer à l'apprentissage d'un de ces deux segments. Mais la réalité est tout autre: parfois ces approximantes centrales rétroflexes sont prononcées telles quelles en québécois, et parfois elles se voient transformées en vibrantes. Il est vrai qu'il y a des dialectes, notamment dans l'est et le nord-est du Québec, où /ɹ/ devient systématiquement /ʀ/, et où on entendra, par exemple, *tire* francisé à /tæyœ:ʀ/ plutôt qu'à /tæyəɹ/, mais ce ne sont pas les locuteurs de ces régions que les anglophones risquent d'entendre le plus souvent.

Et puis, lesquels de ces *r* sont conservés et lesquels sont modifiés au juste? Dans les lexèmes ordinaires, il y a une forte tendance à changer /ɹ/ à /ʀ/ en début de syllabe, comme dans *brake/break*, *rock*,

*drum*, *ratchet*, *trip*, *track*, *thrill*, *strip*, etc., et à le prononcer à l'anglaise en fin de syllabe, comme dans *starter*, *lighter*, *zipper*, *plaster*, *poster*, et tous les autres mots en *-er*. Par contre, dans les noms propres, il me semble qu'à peu près tous /ɹ/ se prononcent tels quels, p. ex. les *Rangers*, *les Bruins*, *les Red Wings*, *Rochester*, *Grand Falls*, *Trenton*, etc. Il est important de noter que ce ne sont pas seulement les gens bilingues qui prononcent de cette façon: tous les locuteurs des régions dialectales qui ne sont pas historiquement grasseyantes semblent avoir le segment /ɹ/ dans leur inventaire phonologique.

On ne peut évidemment pas parler des mots anglais en québécois sans faire allusion au cas classique de l'assibilation dont l'opération, aux yeux — ou plutôt aux oreilles — d'un anglophone, doit sembler tout à fait capricieuse. Pourquoi les /ti/ et /di/ de sa langue maternelle deviendraient-ils /t si/ et /d zi/ dans des formes comme *steam*, *tip*, *dispatcher*, *stew* ou *tune-up*, mais non dans *meeting*, *deal*, *steering*, *building* ou *T-bone*? Pourquoi y aurait-il une affriquée dans *Cincinnati* et non dans *Ypsilanti*, dans *Tiffany* et non dans *Timmins*? *Go figure*!

Enfin, il y a des phénomènes qui, bien que foncièrement réguliers, doivent quand même constituer des sources de mystification et de frustration pour les anglophones. Il ne doit pas être facile, par exemple, de s'habituer au fait que la voyelle finale de tous les mots qui se terminent en *-y*, et qu'ils identifient au timbre vocalique de *be*, *see*, *tea* même si elle est moins longue et souvent plus relâchée qu'un /i/ accentué,[3] se voit systématiquement transformée en mi-ouverte de sorte qu'il faille prononcer *brandy*, *caddy*, *party*, *Johnny*, *Calgary*, *Presley*, etc. avec un /e/ final.[4]

[3] Comme le fait remarquer Bronstein, "when the final syllable is unstressed, as in *city* or *pity*, most speakers seem to use a shortened, lowered, and lax form of /i/ [...]" (1960:147).

[4] Les Français prononcent bien les -y anglais en /i/, mais ceci tient sans doute du fait qu'ils adaptent généralement les mots d'emprunt beaucoup plus à partir de leur forme orthographique que de leur structure phonologique, comme le démontrent des formes comme /büfalo/ pour *Buffalo*, /šatɛʀtɔn/ pour *chatterton*, /šwiŋgɔm/ pour *chewing-gum*, etc.

## 4. Conclusion

En somme, on aurait peut-être tendance à croire de prime abord que, en ce qui a trait à la prononciation, les anglophones nord-américains sont avantagés par rapport aux autres apprenants du français langue seconde du fait qu'il y ait tellement de lexèmes et surtout de noms propres qui leur soient familiers. Or je crois avoir réussi à démontrer que c'est plutôt le contraire qui se produit car le réflexe de prononcer toutes ces formes à l'anglaise paraît difficile à réprimer. À cela vient s'ajouter le problème de l'imprévisibilité qui les empêche plus souvent qu'autrement d'avoir recours à des règles générales d'adaptation des mots anglais en français, surtout lorsqu'il s'agit de noms propres. Reste à savoir si c'est le phénomène décrit ci-haut qui est vraiment singulier et exceptionnel, ou bien si ce n'est pas plutôt le peu d'attention qu'on y a prêté qui constitue la véritable anomalie. S'il s'avère que ce genre de situation est courante et répandue dans les langues du monde, on ne devrait pas tarder à le savoir.

## Références

Bronstein, Arthur J. 1960. *The pronunciation of American English: an introduction to phonetics*. New York: Appleton-Century-Crofts.

Dell, François. 1973. *Les règles et les sons: introduction à la phonologie générative*. Paris: Hermann.

Fouché, Pierre. 1959. *Traité de prononciation française*. Paris: Klincksieck.

Grammont, Maurice. 1914. *Traité pratique de prononciation française*. Paris: Delagrave.

Paradis Carole & Darlene LaCharité. 1997. "Preservation and minimality in loanword adaptation". *Journal of Linguistics* 33:1 [à paraître].

Picard, Marc. 1979. "The syllable boundary in generative phonology". *Revue canadienne de linguistique* 24:119-131.

Picard, Marc. 1983. "La productivité des règles phonologiques et les emprunts de l'anglais en québécois". *Revue de l'Association québécoise de linguistique* 3:97-108.

Picard, Marc. 1996. *The Comparative Phonetics of English and*

*French*. Montréal: Concordia University.
Picard, Marc & Janet Nicol. 1982. "Vers un modèle concret de la phonologie des emprunts". *Revue canadienne de linguistique* 27:156-169.
Schane, Sanford. 1968. *French Phonology and Morphology*. Cambridge, MA: MIT Press.
Tranel, Bernard. 1987. *The Sounds of French*. Cambridge: Cambridge University Press.

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 253-264.

# Phonologie et morphologie de la langue emprunteuse: intégration des emprunts français en kinyarwanda*

*Yvan Rose*
Université McGill

## 1. Introduction

Cet article a pour but de montrer que, premièrement, une analyse dans le cadre de la *théorie des contraintes et stratégies de réparation* (TCSR; Paradis 1988a,b) permet de faire d'importantes prédictions quant à l'adaptation phonologique d'emprunts. Plus précisément, nous verrons que les voyelles orales des emprunts français qui sont mal formées en kinyarwanda sont adaptées de manière prévisible et gouvernée par les principes de la TCSR.[1] Deuxièmement, nous verrons que les emprunts français étudiés ici font partie intégrante du lexique du kinyarwanda.

Avant d'aborder cette analyse, je me dois de préciser les préalables qui ont guidé l'étude des emprunts français en kinyarwanda. Dans la lignée de Paradis (ce volume) et Paradis et LaCharité (1996, 1997), je soutiens l'idée que:

> — Les emprunts sont introduits principalement par des bilingues, qui ont accès au code phonologique de la langue source, une hypothèse appuyée par de nombreuses études sociolinguistiques comme Haugen (1950), Mougeon, Beniak et Valois (1985), Poplack, Sankoff et Miller (1988).
> — Les patrons phonologiques d'emprunts sont imposés à l'ensemble de la communauté par les bilingues, surtout quand une communauté atteint un stade de bilinguisme moyen ou avancé.
> — Un emprunt est un mot simple de la langue source, ou un composé fonctionnant comme un mot simple, qui se conforme

* Je remercie Julie Auger, Denis Dumas, Heather Goad, Éliane Lebel, Carole Paradis, Glyne Piggott, Shana Poplack et Gillian Sankoff pour leur questions et commentaires lors de la présentation orale de cette recherche, au cours du colloque *Explorations du lexique*, qui s'est tenu dans le cadre du congrès annuel 1996 de l'ACFAS. L'auteur de cet article a bénéficié d'une bourse de maîtrise du Fonds FCAR (concours B1-11B) et d'une bourse de doctorat du CRSH #752-95-1415.

[1] Pour des raisons d'espace, l'analyse des voyelles nasales et l'adaptation des suites segmentales françaises mal formées en kinyarwanda ne pourront être abordées dans cet article. Pour plus de détails au sujet de ces adaptations, voir Rose (1995).

(au moins) aux contraintes phonologiques les plus externes de la périphérie de la langue emprunteuse, et qui est incorporé dans le discours de la langue emprunteuse.
— La forme empruntée n'est pas l'output phonétique de la langue source, mais son output lexical ou syntaxique.
— Les emprunteurs éliminent des emprunts l'information qu'ils perçoivent comme redondante du point de vue de la langue emprunteuse à moins que cette information soit perçue comme distinctive dans la langue source. Les combinaisons de traits distinctives de la langue source sont toutes systématiquement introduites dans la langue emprunteuse.

L'article se divise comme suit: le cadre théorique et la structure des langues étudiées sont brièvement présentés en 2. Nous verrons l'analyse de l'adaptation des voyelles orales en 3. Les exemples d'intégration lexicale des emprunts sont présentés en 4.

Le corpus étudié comprend 756 emprunts français en kinyarwanda vérifiés auprès de trois informateurs. Les statistiques en (1) montrent que les structures mal formées en kinyarwanda sont adaptées dans 98% des cas.

(1) Statistiques sur les emprunts français en kinyarwanda

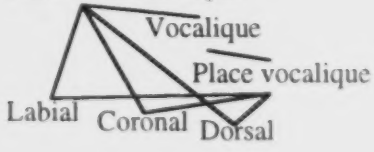
| | Adaptations syllabiques | Adaptations segmentales | Total | % |
|---|---|---|---|---|
| Cas analysés | 2 420 | 1 936 | 4 356 | |
| Adaptations | 2 335 | 1 933 | 4 268 | 98% |
| Non-adaptations | 26 | 0 | 26 | 0,6% |
| Élisions | 59 | 3 | 62 | 1,4% |

## 2. Cadre théorique et structure des langues étudiées

### 2.1 Représentations phonologiques

J'utilise un modèle de représentation basé sur Clements et Hume (1995). La hiérarchie des articulateurs consonantiques et vocaliques qu'ils proposent est présentée en (2).

(2)



## 2.2 Processus phonologiques

Comme il a été mentionné dans l'introduction, je situe mon analyse dans le cadre de la TCSR. La TCSR est une théorie de contraintes en phonologie qui pose que les langues sont gouvernées par des contraintes universelles (principes) et non universelles (réglages paramétriques). Il arrive que les contraintes d'une langue soient violées. Dans ce cas, la violation est réparée par une stratégie de réparation, un processus phonologique qui vise à préserver la contrainte violée. Les stratégies de réparation s'appliquent en fonction des principes en (3), (4) et (5).[2]

(3) Principe de minimalité: Une stratégie de réparation
a) doit s'appliquer au niveau le plus bas auquel fait référence la contrainte violée;
b) doit impliquer le moins de stratégies (d'étapes) possible.

(4) Principe de préservation:
L'information segmentale est préservée maximalement, à l'intérieur des limites du principe du seuil de tolérance.

(5) Principe du seuil de tolérance:
a) Toutes les langues imposent une limite à la préservation segmentale;
b) cette limite est fixée à deux étapes (deux stratégies de réparation) à l'intérieur d'un domaine de contrainte.

Un aperçu de la TCSR est présenté en (6).

(6) TCSR:

| | |
|---|---|
| Violation de contrainte | Les stratégies de réparation sont gouvernées par: |
| ↓ | • Principe de préservation |
| Stratégie de réparation | • Principe du seuil de tolérance<br>• Principe de minimalité } basés sur la hiérarchie des<br>• Convention de préséance } niveaux phonologiques |

## 2.3 Structure des langues étudiées

Les systèmes phonologiques des langues étudiées dans cet article sont présentés respectivement en (7) et (8).

---

[2] Pour une définition plus explicite de ces principes et de la TCSR en général, voir, entre autres, Paradis et LaCharité (1996, 1997) et Paradis (ce volume).

(7) Système phonologique du français (Casagrande 1984)

| | Labial | Coronal | | Dorsal |
|---|---|---|---|---|
| occl. | p / b | t / d | | k / g |
| fric. | f / v | s / z | ʃ / ʒ | |
| nas. | m | n | ɲ | |
| liq. | | l / ʁ | | |
| glide | | | j / ɥ | w |

| | | |
|---|---|---|
| i | y | u |
| e | ø | o |
| ɛ / ɛ̃ | œ / œ̃ | ɔ / ɔ̃ |
| | (ə) | |
| | a / ɑ̃ | |

(8) Système phonologique du kinyarwanda (Kimenyi 1979)

| | Labial | Coronal | | Dorsal | Laryn. |
|---|---|---|---|---|---|
| occl. | p | t / d | | k / g | |
| fric. | β / f / v | s / z | ʃ / ʒ / ç | | h |
| affr. | | ts | tʃ | | |
| nas. | m | n | ɲ | ŋ | |
| liq. | | r | | | |
| glide | | | j | w | |

| | |
|---|---|
| i / ii | u / uu |
| e / ee | o / oo |
| a / aa | |

On peut noter, en comparant les systèmes en (7) et (8) que, contrairement au kinyarwanda, le français comprend des voyelles nasales, des voyelles labiales coronales comme *y* et *ø*, et affiche un contraste voyelles tendues / non tendues, par exemple, entre *e* et *ɛ*. Toutes les voyelles du français autres que *i, e, a, o, u*, soit les voyelles en (9), doivent donc être adaptées dans les emprunts français en kinyarwanda.

(9) Voyelles françaises problématiques pour le kinyarwanda[3]

a) voyelles coronales labiales: *y, ø* et *œ*
b) voyelles non tendues: *ɛ* et *ɔ*
c) voyelles nasales: *ɑ̃, ɛ̃, ɔ̃* et *œ̃*

## 3. Adaptation des voyelles orales

### 3.1 Adaptation des voyelles labiales coronales *y, ø* et *œ*

Contrairement au français, le kinyarwanda ne permet pas la combinaison des articulateurs Labial et Coronal en forme sous-

[3] Comme le statut du chva français (*ə*) demeure incertain à bien des égards et que cette voyelle présente un comportement phonologique très irrégulier, aucune analyse de son adaptation n'a été effectuée. Voir Rose (1995) pour plus de détails.

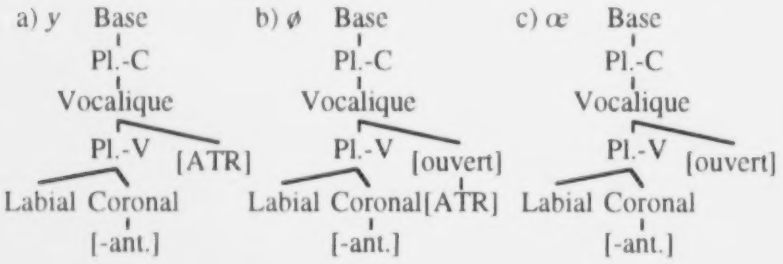
jacente (FSJ). Cette contrainte du kinyarwanda est exprimée formellement par le réglage paramétrique négatif présenté en (10).

(10) Paramètre:

| | | |
|---|---|---|
| Labial-Coronal? | français: | oui |
| | kinyarwanda: | **non** (**y*, **ø*, **œ*) |

Comme on peut le voir en (11), la FSJ des voyelles *y*, *ø* et *œ* comporte simultanément les articulateurs Labial et Coronal, ce qui viole la contrainte (10).

(11) FSJ de *y*, *ø* et *œ* (représentation simplifiée)[4]




Des exemples d'emprunts contenant *y*, *ø* ou *œ* sont présentés en (12). Comme on peut le constater, ces voyelles sont généralement adaptées en voyelles non arrondies en kinyarwanda.

(12) Adaptation des voyelles *y*, *ø* et *œ*

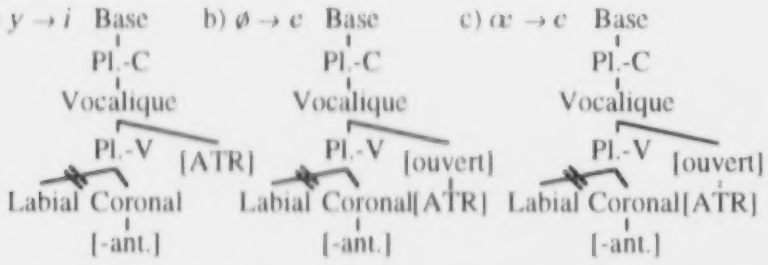
| | *français* | | *kinyarwanda* |
|---|---|---|---|
| a) Voyelle *y* | ambulance | [ɑ̃by̲lɑ̃s] | → [a$^{m}$bi̲raa$^{n}$si] |
| | député | [depy̲te] | → [depi̲te] |
| | porte-plume | [pɔʁtəply̲m] | → [porotopuri̲mi] |
| | reçu | [ʁɔsy̲] | → [rosi̲] |

[4] Les voyelles hautes (*i*, *y* et *u*) sont représentées sans le trait [ouvert] sous Vocalique, contrairement aux voyelles non hautes, qui doivent être spécifiées pour [ouvert] (voir Goad 1993).

| | | | |
|---|---|---|---|
| b) Voyelle ø | eucalyptus | [økaliptys] | → [ɕkariβutusi] |
| | lieutenant | [ljøtnɑ̃] | → [rijɕtona] |
| | mitrailleuse | [mitʁajøz] | → [mitarajɕɕze][5] |
| | poudreuse | [pudʁøz] | → [puderɕɕze] |
| c) Voyelle œ | ambassadeur | [ɑ̃basadœʁ] | → [aᵐbasadɕɕri] |
| | convoyeur | [kɔ̃vwajœʁ] | → [kʷooᵐvuwajɕɕri] |
| | docteur | [dɔktœʁ] | → [dojitɕɕri] |
| | manoeuvre | [manœvʁ] | → [manɕɕvure] |

L'adaptation générale des voyelles *y*, ø et œ consiste donc à dissocier l'articulateur Labial de la voyelle à adapter. C'est ce que l'on peut voir dans les processus en (13a, b et c). L'élision de l'articulateur Labial produit une voyelle coronale bien formée en kinyarwanda.

(13) Labial-Coronal → Coronal




Le choix de dissocier Labial plutôt que Coronal ne peut pas être prédit phonologiquement, étant donné que ces adaptations sont toutes deux minimales, selon le principe de minimalité en (3). Ce choix peut toutefois être relié à une convention sociale dans l'adaptation des emprunts (voir Paradis et LaCharité 1996, 1997 à ce sujet).

[5] De prime abord, aucun indice, que ce soit la qualité des voyelles allongées, les consonnes qui apparaissent dans l'environnement de ces voyelles, ou la position de ces voyelles à l'intérieur des mots, ne permet de prédire leur allongement dans les emprunts en kinyarwanda. De plus amples recherches devront toutefois être effectuées.

### 3.2 Adaptation des voyelles non tendues ɛ et ɔ

Les voyelles françaises non tendues (non [ATR]) ɛ et ɔ violent la contrainte (14) qui interdit les voyelles non tendues en kinyarwanda.

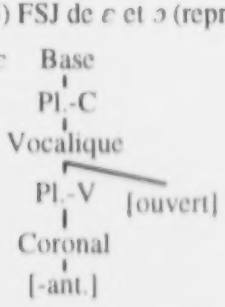
(14) Paramètre:

| | | |
|---|---|---|
| Voyelles non tendues? | français: | oui |
| | kinyarwanda: | **non** (*ɛ, *ɔ) |

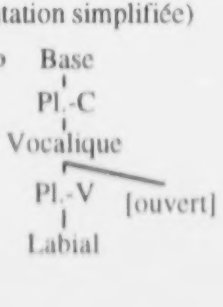
Ces voyelles doivent donc être adaptées dans les emprunts français en kinyarwanda. On peut voir en (15a et b) que le trait [ATR] est absent de la FSJ des voyelles ɛ et ɔ, ce qui viole la contrainte (14).

(15) FSJ de ɛ et ɔ (représentation simplifiée)

a) ɛ




b) ɔ




Des exemples d'emprunts qui contiennent les voyelles ɛ et ɔ sont présentés en (16). Comme on peut s'y attendre, ces deux voyelles sont adaptées en voyelles tendues en kinyarwanda.

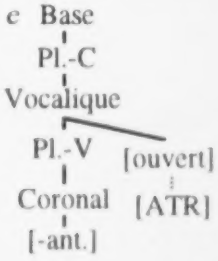
(16) Adaptation des voyelles ɛ et ɔ

| | *français* | | | *kinyarwanda* |
|---|---|---|---|---|
| a) Voyelle ɛ | arrêt | [aʁɛ] | → | [arɛ] |
| | boulette | [bulɛt] | → | [βurɛti] |
| | certificat | [sɛʁtifika] | → | [sɛretifika] |
| | électricité | [elɛktʁisite] | → | [erɛjitirisite] |
| | sifflet | [siflɛ] | → | [siifurɛ] |
| b) Voyelle ɔ | démocratie | [demɔkʁasi] | → | [demokarasi] |
| | colonie | [kɔlɔni] | → | [kʷoroni] |
| | diocèse | [diɔsɛz] | → | [dijoseezi] |
| | laboratoire | [labɔʁatwaʁ] | → | [raβoratkʷaari] |
| | limonade | [limɔnad] | → | [rimonaadi] |

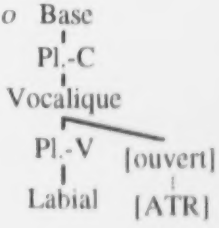
Étant donné que le seul moyen d'adapter une voyelle non tendue est d'insérer le trait [ATR] dans sa représentation, c'est cette stratégie qui est appliquée. Le processus d'insertion du trait [ATR] est représenté en (17a et b).

(17) Ø → [ATR] (insertion du trait [ATR])

a) ɛ → e




b) ɔ → o




Les adaptations que nous avons vues dans cette section s'appliquent avec un minimum d'étapes, et au niveau le plus bas auquel font référence les contraintes (10) et (14), en accord avec le principe de minimalité en (3). La dissociation de Labial des voyelles labiales coronales en (13) viole le principe de préservation en (4). Mais comme cette élision est directement causée par le contexte phonologique, elle demeure tout à fait prévisible; aucune insertion ne pourrait rendre la combinaison Labial-Coronal bien formée en kinyarwanda. Finalement, les adaptations vues ici s'appliquent toutes en deux étapes ou moins, ce qui est conforme au principe du seuil de tolérance en (5).

## 4. Intégration lexicale

En plus d'être adaptés dans la presque totalité des cas (les non-adaptations ne comptant que pour 0,6% des cas; voir les statistiques en (1)), les emprunts français qui ont été importés au Rwanda sont aussi parfaitement intégrés morphologiquement.

Il est tout d'abord nécessaire de préciser que le kinyarwanda est une langue à marqueurs de classe. Par exemple, la classe de noms [+animé] est marquée par le préfixe *mu-* au singulier et le préfixe *ba-* au pluriel. On peut voir en (18) que les noms français [+animé] qui sont adaptés en kinyarwanda reçoivent aussi les préfixes marqueurs de classe *mu-* et *ba-*. On voit par exemple que le mot *candidat* devient [(u)mukaaⁿdida] au singulier et [(a)βakaaⁿdida] au pluriel.

(18) Noms français [+animé] avec préfixes marqueurs de classe

| *français* | | *kinyarwanda* | |
|---|---|---|---|
| | | *Singulier* | *Pluriel* |
| candidat | [kɔ̃dida] | [(u)mu-kaa$^{n}$dida] | [(a)βa-kaa$^{n}$dida][6] |
| détaillant | [detajɑ̃] | [(u)mu-detaja] | [(a)βa-detaja] |
| gendarme | [ʒɑ̃daʁm] | [(u)mu-ʒaa$^{n}$darume] | [(a)βa-ʒaa$^{n}$darume] |
| journalier | [ʒuʁnalje] | [(u)mu-ʒurunarije] | [(a)βa-ʒurunarije] |
| juge | [ʒyʒ] | [(u)mu-ʒuuʒi] | [(a)βa-ʒuuʒi] |
| lieutenant | [ljøtnɑ̃] | [(u)mu-rijetona] | [(a)βa-rijetona] |
| maçon | [masɔ̃] | [(u)mu-maso] | [(a)βa-maso] |
| pasteur | [pastœʁ] | [(u)mu-paasitoori] | [(a)βa-paasitoori] |
| photographe | [fɔtɔgʁaf] | [(u)mu-fotogaraafe] | [(a)βa-fotogaraafe] |
| préfet | [pʁefɛ] | [(u)mu-perefe] | [(a)βa-perefe] |

Les affixes du kinyarwanda affectent aussi les verbes du français qui sont adaptés dans cette langue. Je vais ici discuter de l'infinitif, de l'impératif et du perfectif en kinyarwanda. À l'infinitif, les verbes du kinyarwanda reçoivent le préfixe *g$^{w}$u-* et se terminent par le suffixe *-a*. Les verbes à l'impératif ont la forme de l'infinitif mais sans le préfixe *g$^{w}$u-*. Quant aux verbes perfectifs, ils reçoivent le préfixe *jara-* et le suffixe *-je*. On peut observer en (19) que les verbes français empruntés se conforment en tous points aux patrons morphologiques du kinyarwanda. Par exemple, le verbe français *chauffer* devient en kinyarwanda [g$^{w}$u-ʃof-a] à l'infinitif, [ʃof-a] à l'impératif et [jara-ʃof-je] au perfectif.

[6] Les voyelles entre parenthèses ((*u*) et (*a*)) sont des pré-préfixes (augments) qui consistent en la réduplication de la voyelle du préfixe dans les contextes où le mot n'est précédé d'aucune autre voyelle.

(19) Verbes français adaptés en kinyarwanda

| *français* | | *kinyarwanda* | | |
|---|---|---|---|---|
| | | *Infinitif* | *Impératif* | *Perfectif* |
| chauffer | [ʃofe] | [gʷu-ʃof-a] | [ʃof-a] | [jara-ʃof-je] |
| baptiser | [batize] | [kʷu-βatiz-a][7] | [βatiz-a] | [jara-βatiʒ-e][8] |
| chasser | [ʃase] | [gʷu-ʃas-a] | [ʃas-a] | [jara-ʃaʃ-e] |
| couper | [kupe] | [gʷu-kʷup-a] | [kʷup-a] | [jara-kʷup-je] |
| payer | [peje] | [gʷu-pej-a] | [pej-a] | [jara-pej-e] |
| photographier | [fɔtɔgʁafje] | [gʷu-foto-ra][9] | [foto-ra] | [jara-foto-je] |
| signer | [siɲe] | [gʷu-siɲ-a] | [siɲ-a] | [jara-siɲ-e] |

Nous pouvons aussi observer que, en accord avec les présupposés théoriques mentionnés en introduction, les verbes adaptés en kinyarwanda ne contiennent que le radical du verbe français; aucun suffixe flexionnel français ne se retrouve dans les verbes adaptés en kinyarwanda. Ceci supporte l'hypothèse que les locuteurs emprunteurs sont réellement des bilingues qui ont accès au code de la langue source (ici, le français) et que la forme empruntée constitue bien une entrée lexicale (et non une simple forme phonétique).

De plus, il serait intéressant de vérifier de quelle manière le contexte d'emprunt des mots français en kinyarwanda diffère de celui des mots anglais en français cadjin. Dubois et Sankoff (ce volume) rapportent que certains mots anglais en français cadjin ne sont adaptés ni phonologiquement, ni morphologiquement. Par exemple, le participe passé de la phrase cadjin "J'ai *draw* mon security check" ne porte aucun suffixe marqueur de temps. Une comparaison entre les contextes sociaux d'importation d'emprunts en kinyarwanda et en cadjin ainsi qu'une étude plus approfondie de la grammaire de ces deux langues pourraient sans doute fournir plus de précisions sur les raisons de ces résultats contradictoires.

---

[7] La prononciation [kʷu-] du préfixe /gʷu-/ est causée par une règle de dissimilation en kinyarwanda: la consonne du préfixe est toujours dévoisée lorsque la première consonne du radical est voisée et vice versa (voir aussi les autres exemples dans le tableau).

[8] Le /z/ du radical de *baptiser* est fusionné au /j/ du suffixe du perfectif /-je/. C'est pourquoi il se prononce [ʒ]. Cette fusion se produit dans tous les cas où le radical se termine par une consonne coronale.

[9] Il semble que le *r* soit ici inséré pour briser une séquence VV interdite en kinyarwanda. Cette hypothèse doit toutefois être étudiée plus en détail.

## 5. Conclusion

En résumé, nous avons vu que, d'un point de vue phonologique, les segments des emprunts français en kinyarwanda sont adaptés avec une très grande régularité. Toutes les adaptations phonologiques s'appliquent de façon minimale, c'est-à-dire au niveau de la contrainte violée, et en un minimum d'étapes, conformément au principe de minimalité. De plus, nous avons vu que les segments sont adaptés avec une préservation maximale de l'information segmentale, en accord avec le principe de préservation. Les élisions d'information segmentale sont toutes causées par le contexte phonologique, qui exige la dissociation plutôt que l'insertion d'un trait. Ces résultats correspondent à ceux des analyses d'emprunts français en peul (voir, entre autres, Lebel 1994, Paradis et LaCharité 1996, 1997, et Paradis, ce volume) et d'emprunts anglais en français québécois (voir, entre autres, Paradis, Lebel et LaCharité 1993 et Paradis et Lebel, ce volume).

D'un point de vue morphologique, nous avons vu que les emprunts français font partie intégrante du lexique du kinyarwanda. Ils sont dérivés avec les mêmes affixes et marqueurs de classes que les mots natifs et se conforment en tous points aux temps et modes verbaux du kinyarwanda. On ne peut donc pas parler de transfert de code ni de lexique parallèle ou étranger au kinyarwanda, mais bien d'entrées lexicales totalement intégrées à la langue emprunteuse.

Une question majeure ressortissant de cette étude est que les résultats obtenus ici sont contradictoires à ceux obtenus par Dubois et Sankoff (ce volume) dans leur étude d'emprunts anglais en français cadjin. Un étude plus précise du contexte d'emprunt entre les langues impliquées pourrait éventuellement éclaircir cette question et, peut-être, préciser davantage la notion d'emprunt ainsi que les conditions — linguistiques et sociales — qui régissent son intégration dans la langue emprunteuse.

## Bibliographie

Casagrande, Jean (1984). *The Sound System of French*. Washington, DC: Georgetown University Press.

Clements, George N. et Elizabeth V. Hume (1995). The Internal Organization of Speech Sounds. Dans John A. Goldsmith (réd.) *The Handbook of Phonological Theory*. Cambridge, Mass.: Blackwell, 245-306.

Dubois, Sylvie et David Sankoff (ce volume). L'absence de flexion sur les emprunts à l'anglais dans le français cadjin. Dans Julie Auger et Yvan Rose (réd.) *Explorations du lexique*. Québec: CIRAL.

Goad, Heather (1993). *On the Configuration of Height Features*. Thèse de doctorat. University of Southern California at Los Angeles.

Haugen, Einar (1950). The Analysis of Linguistic Borrowings. *Language* 26: 210-231.

Kimenyi, Alexandre (1979). *Studies in Kinyarwanda and Bantu Phonology*. Current Inquiry into Language and Linguistics 33. Edmonton, Illinois: Linguistic Research.

Lebel, Caroline (1994). *Le rôle des contraintes et stratégies de réparation dans l'adaptation phonologique d'emprunts français en peul*. Mémoire de maîtrise. Université Laval.

Mougeon, Raymond, Édouard Beniak et Daniel Valois (1985). *Variation in the Phonological Integration of Loanwords in a Bilingual Speech Community*. Toronto: Center for Franco-Ontarian Studies.

Paradis, Carole (1988a). On Constraints and Repair Strategies. *The Linguistic Review* 6: 71-97.

___ (1988b). Towards a Theory of Constraint Violations. *McGill Working Papers in Linguistics* 5(1): 1-43.

___ (ce volume). Préservation phonémique et seuil de tolérance dans l'adaptation d'emprunts. Dans Julie Auger et Yvan Rose (réd.) *Explorations du lexique*. Québec: CIRAL.

Paradis, Carole et Darlene LaCharité (1996). Saving and Cost in French Loanword Adaptation in Fula: Predictions of the TCRS Loanword Model. *McGill Working Papers in Linguistics* 11: 46-84.

___ (1997). Preservation and Minimality in Loanword Adaptation. *Journal of Linguistics* 33(1). [sous presse]

Paradis, Carole, Caroline Lebel et Darlene LaCharité (1993). Adaptation d'emprunts: les conditions de la préservation segmentale. Dans Carrie Dyck (réd.) *Actes du Congrès annuel de l'Association canadienne de linguistique*. Toronto: Toronto Working Papers in Linguistics, 461-476.

Paradis, Carole et Éliane Lebel (ce volume). Centre et Périphérie: comportement des segments anglais en français québécois. Dans Julie Auger et Yvan Rose (réd.) *Explorations du lexique*. Québec: CIRAL.

Poplack, Shana, David Sankoff et Christopher Miller (1988). The Social Correlates and Linguistic Processes of Lexical Borrowing and Assimilation. *Linguistics* 26: 47-104.

Rose, Yvan (1995). *Minimalité, préservation et tolérance dans les emprunts français en kinyarwanda*. Mémoire de maîtrise. Université Laval.

# Section 5

*Lexiques spécialisés*

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 267-280.

# Quelques aspects particuliers du traitement prédictionnairique du vocabulaire général d'orientation technique

Hélène Cajolet-Laganière et Normand Maillet
Université de Sherbrooke

Aujourd'hui, c'est devenu un lieu commun d'affirmer que les technologies caractériseront l'histoire du XX$^{e}$ siècle. En effet, les progrès fulgurants des moyens de transport, des télécommunications, le développement rapide des procédés, machines et techniques de fabrication ont modifié considérablement les usages sociaux, culturels, politiques et économiques. Tous ces bouleversements engendrent d'importantes conséquences linguistiques.

## Importance des termes techniques

En effet, les récentes études démontrent que les textes techniques et scientifiques constituent plus de 60 % de toute la production écrite en français. La quantité des termes techniques et scientifiques s'accroît de manière significative, au point d'influencer la structure des nomenclatures des dictionnaires de langue générale. La proportion des termes spécialisés dans ces ouvrages lexicographiques augmente d'une édition à l'autre, représentant aujourd'hui, selon les lexicographes, plus de 40 % de la nomenclature d'un dictionnaire usuel (Boulanger 1989). La question du choix des mots et des sens techniques, des cooccurrents, des mots en concurrence (synonymes), des exemples et des citations prend ainsi une grande importance dans l'établissement de la nomenclature et dans le traitement dictionnairique de ces données lexicales. En outre, l'intérêt accordé aux termes dans les ouvrages lexicographiques généraux a amené le développement d'un système de marques socioprofessionnelles qui accompagnent et caractérisent les emplois des mots techniques ou scientifiques. On note cependant que la délimitation entre langue commune et langue technique demeure en grande partie subjective. La variation

de la qualification et du marquage des termes techniques d'un lexicographe à un autre en est certes une preuve éloquente. Notre recherche se situe dans ces axes de réflexion et d'analyse, soit la caractérisation lexicale de la langue technique québécoise et son traitement lexicographique dans le cadre d'un dictionnaire général du français québécois.

**Définition de la langue technique**

Aux fins de notre recherche, nous appelons « langue technique » la langue propre aux spécialités quand on les considère en elles-mêmes, surtout au stade des manipulations, des applications pratiques, de la fabrication ou de la production. Nous la distinguons de la langue scientifique qui, pour nous, est plutôt liée aux opérations intellectuelles supposant toute démarche d'analyse, de recherche, d'induction ou de déduction. Ces distinctions nous sont apparues plus fonctionnelles pour établir notre corpus, puisqu'il y a beaucoup d'interrelations (sciences et techniques) dans le développement des technologies contemporaines. Les domaines eux-mêmes sont soumis à ces interférences et ne constituent donc pas, à eux seuls, un critère sûr de délimitation. On remarque d'ailleurs que dans les études antérieures, les chercheurs ont parlé de langue scientifique en incluant parfois certaines techniques, ou de langue technoscientifique, laissant entendre ainsi que les deux possèdent les mêmes caractéristiques. Cependant, la langue technique n'a pas été décrite en elle-même et les domaines analysés jusqu'à maintenant ont plutôt été scientifiques : mathématiques, chimie, physique, sciences naturelles, médecine, astronautique, etc. De plus, les recherches ont porté principalement sur les langues anglaise et allemande; très peu ont été faites sur le français, encore moins évidemment sur le français québécois (Gambier 1995). C'est pourquoi, dans le cadre des recherches du CATIFQ (Centre d'analyse et de traitement informatique du français québécois) portant sur la caractérisation et la description du vocabulaire technique québécois, nous avons décidé d'effectuer deux études distinctes, l'une sur la langue technique et l'autre sur la langue

scientifique, faisant appel à deux corpus de textes différents. Une étude comparative des deux corpus nous permettra de cerner, selon le cas, les spécificités de ces deux discours.

**Corpus d'analyse**

Pour la langue technique québécoise, notre corpus d'analyse est constitué de 100 tranches d'égale longueur, soit de 2500 mots, comptés selon une méthodologie déterminée qui assure l'uniformité de la taille de celles-ci. Notre analyse porte donc sur un corpus de 250 000 mots. Les textes que nous avons utilisés pour extraire les tranches de notre corpus d'analyse ont été sélectionnés à l'aide de critères précis. (Voir ci-dessous les domaines sélectionnés ainsi que la provenance et la typologie des textes répertoriés. En outre, nous n'avons retenu que des rédacteurs et rédactrices de langue maternelle française, ayant fait des études primaires et secondaires au Québec et, le cas échéant, des études collégiales ou universitaires en français au Québec. Enfin, comme il s'agit d'une étude synchronique, tous les textes ont été rédigés après 1980. Ces critères ont été élaborés pour répondre le mieux possible aux exigences d'homogénéité spatiale, temporelle et rédactionnelle, pour rendre compte de la variation communicationnelle et terminologique à l'intérieur de domaines représentatifs de l'activité technique québécoise.

## CORPUS

| Choix de dix domaines représentatifs de l'activité technique québécoise | Typologie des textes reflétant diverses situations de communication |
|---|---|
| ALUMINIUM - Alcan<br><br>MINES - Ministère Énergie Ressources<br>- Association minière du Québec<br>- Centre de recherche minier<br><br>TÉLÉCOMMUNICATIONS<br>- Bell Canada<br>- Québec téléphone<br><br>INFORMATIQUE<br>- CRIM - MICT<br>- Ministère des Transports<br>- Ville de Montréal<br><br>ENVIRONNEMENT<br>- Hydro-Québec<br><br>HYDRO-ÉLECTRICITÉ<br>- Hydro-Québec<br><br>PÂTES ET PAPIERS<br>- Linguatec<br>- Stone-Consolidated<br><br>TRANSPORT<br>- Ministère des Transports<br><br>DIVERS 1 (construction, essai et matériaux, béton, etc.)<br>- Domatec<br>- Ministère des Transports<br><br>DIVERS 2 (hydraulique, maintenance, gestion de projets, etc.)<br>- APIBQ<br>- Cégep de Trois-Rivières<br>- Ministère des Transports<br>- MICT | . Rapports techniques<br>. Rapports d'évaluation<br>. Rapports d'avant-projet<br>. Manuels de formation<br>. Monographie<br>. Guides d'entretien<br>. Guides de construction<br>. Cahier de charges<br>. Études techniques<br>. Enquêtes<br>. Normes<br>. Codes de procédures |

**Méthodologie de traitement du corpus**

La préparation des tranches du corpus en vue de l'analyse statistique a nécessité plusieurs étapes de traitement et a été faite conformément aux normes de dépouillement et de traitement en usage au CATIFQ. Cela touche le tirage au sort des tranches, le comptage des mots, l'indexation, la lemmatisation, le traitement des syntagmes, etc. Pour ces divers aspects, nous renvoyons le lecteur à l'article « Caractérisation des textes techniques québécois », publié dans le numéro 47 de la revue *Présence francophone*, novembre 1995.

Tout le travail de dépouillement, d'indexation et de lemmatisation étant terminé, nous disposons actuellement d'un index de toutes les données de notre corpus. Il contient environ 9000 mots simples et 6000 syntagmes lexicalisés. Cela nous permettra de tirer un certain nombre de conclusions sur la caractérisation de la langue technique québécoise. Par ailleurs, l'objet de cet article porte sur l'établissement d'une nomenclature de termes et de sens techniques québécois et sur le traitement de celle-ci dans le cadre d'un dictionnaire général du français québécois. En ce qui a trait à la nomenclature, nous nous intéressons aux questions relatives au choix des mots simples et composés (groupes complexes, syntagmes), des sens et des sous-sens, des mots en concurrence (synonymes), des cooccurrents, des citations, etc. En ce qui a trait au traitement, nous nous interrogeons quant au nombre de domaines et de sous-domaines, d'emplois spécialisés et de renseignements encyclopédiques que l'on doit retenir.

L'espace étant limité, nous illustrerons notre propos à l'aide du mot « coulée », choisi parmi les 100 substantifs les plus fréquents. Il a, dans notre corpus, la fréquence 22 et il est présent dans quatre domaines. Selon nos critères de fréquence, de dispersion et d'attestation dans les principaux dictionnaires généraux, il mériterait une entrée dans un dictionnaire général du français québécois. Les cinq tableaux présentés en annexe illustrent les diverses informations recueillies sur ce mot dans les corpus et ouvrages lexicographiques retenus aux fins de notre recherche.

**Le tableau I** montre les quatre sens relevés dans le corpus technique québécois à partir de l'analyse des contextes. Nous les avons mis en parallèle avec les sens répertoriés par la BTQ (Banque de terminologie du Québec), qui est censée refléter l'ensemble des acceptions du mot d'un point de vue technique et terminologique. Nous avons noté en gras les sens exclusivement québécois.

On remarque tout d'abord, et c'était prévisible en raison de son objet et de sa taille, que la BTQ nous donne un grand nombre de sens, soit dix sens techniques différents. Il est aussi intéressant de noter que seulement trois sens sont communs aux deux corpus : les sens 1, 2 (en partie) et 3. En outre, on constate que le traitement du sens 2 est beaucoup plus développé dans la BTQ. Par ailleurs, le sens 4 relevé dans notre corpus de textes techniques est absent de la BTQ, et cela est étonnant. Les sens reliés aux domaines de l'industrie du sucre (8, 9 et 10), à la foresterie (5) et à l'eau (6) n'ont pas été relevés dans le corpus technique, sans doute parce que ces domaines ne figurent pas dans les dix domaines techniques sélectionnés aux fins de notre étude. Enfin, nous notons que le sens technique plus général « d'état du métal coulé refroidi » (sens 7) n'est pas attesté dans notre corpus de textes techniques, même si l'on y traite du coulage de l'aluminium. Ces premières observations démontrent la limite et le caractère complémentaire des informations que peuvent fournir un corpus de textes spécialisés, et même une banque de terminologie, pour établir l'ensemble des sens techniques d'un mot. Il apparaît donc important de prendre en compte les données du corpus de textes techniques et celles fournies dans la BTQ.

**Le tableau II**, pour sa part, compare les emplois du mot *coulée* dans le corpus de textes techniques québécois à ceux relevés dans la BDTS (Banque de données textuelles de Sherbrooke), corpus de plus de 5 000 000 de mots répartis en trois blocs principaux de langue générale ou commune : langue littéraire, langue journalistique et langue orale. On note la correspondance des quatre premiers sens. Seuls les sens reliés à l'industrie du sucre s'ajoutent dans la BDTS. Nous l'avons dit, ceci s'explique par le choix restreint des domaines du corpus technique. Les sens attestés dans la BDTS sont

essentiellement les mêmes que ceux déjà répertoriés dans le corpus technique et la BTQ. Il s'agirait donc vraisemblablement des principaux sens techniques usuels à retenir. Bien que les tableaux ne présentent que les sens, il est important de souligner l'intérêt que représente la BDTS pour ce qui est des cooccurrents et des groupes complexes. À titre d'exemple, notons l'expression *coulée de notes de musique*, relevée dans le corpus de langue littéraire et orale. L'analyse de divers types de corpus apparaît donc essentielle pour faire cette collecte de sens et de sous-sens, de cooccurrents, d'expressions et de groupes complexes.

Dans le **tableau III**, nous comparons les emplois du mot *coulée* tirés du corpus de textes techniques et de la BDTS avec ceux consignés dans les quatre principaux dictionnaires québécois : le *Dictionnaire québécois d'aujourd'hui*, le *Dictionnaire du français plus*, le *Dictionnaire général de la langue française au Canada* et le *Glossaire du parler français au Canada*. On constate une correspondance entre les six sens attestés dans les corpus de textes et la présence de trois nouveaux sens dans les dictionnaires (13, 14, et 15). On note donc neuf sens répertoriés au Québec. Ce constat pose l'épineux problème des limites d'espace reliées à l'édition dictionnairique. Peut-on garder ces neufs acceptions? Quand on compare le traitement lexicographique du mot *coulée* d'un dictionnaire à l'autre, on observe que les choix ont été différents et qu'aucun lexicographe n'a retenu tous ces sens. Par exemple, le DQA présente trois sens (2, 3 et 13); le Bélisle en présente six (1, 2, 3, 11, 14, et 15); quant au DFP et au Glossaire, ils donnent chacun trois sens, qui sont différents du DQA et différents dans l'un et dans l'autre (DFP : 2, 4, et 14); (Glossaire : 3, 11 et 12). Cela illustre l'intérêt de faire la somme des acceptions présentées dans les principaux ouvrages lexicographiques québécois si l'on veut obtenir un éventail complet des sens de ce mot. Cela soulève également la nécessité d'adopter des critères méthodologiques précis quant à la sélection des acceptions.

Dans les deux derniers tableaux, nous avons regroupé l'ensemble des informations recueillies dans les deux corpus

(technique et général) avec celles des quatre dictionnaires québécois et nous les avons comparées, dans un premier temps, avec les données fournies dans quatre dictionnaires français généraux, soit deux dictionnaires usuels (*Petit Robert* et *Petit Larousse illustré*) et deux grands dictionnaires (*Grand Larousse de la langue française* et *Trésor de la langue française)*.

Quand on compare les données colligées à partir des corpus et dictionnaires québécois avec celles consignées dans les deux dictionnaires usuels de France (**Tableau IV**), on constate que seuls les sens 1, 2, 14 et 16 sont communs. Six sens québécois sont absents des dictionnaires français; en revanche, des sens non répertoriés dans les corpus et dictionnaires québécois apparaissent : les sens 17, 18, 19 et 20. Ces derniers sens, même s'ils ne sont pas attestés dans la documentation québécoise analysée, sont probablement connus et en usage au Québec et seraient sans doute attestés si l'on augmentait sensiblement la taille de la BDTS. Ils devraient donc donner lieu à une réflexion au moment de la sélection des acceptions.

Enfin, si l'on considère les grands dictionnaires français (**Tableau V**), huit sens s'ajoutent aux précédents. Ils correspondent à des sens spécialisés reliés à des domaines tels que la pêche, la natation, la navigation, ou encore à des sens figurés. Ainsi, en additionnant tous les sens relevés, on note plus de 25 acceptions du mot *coulée*. Et les synonymes, groupes complexes et cooccurrents n'ont pas encore été pris en compte.

En conclusion, ces observations suscitent un certain nombre de réflexions. D'une part, l'importance de travailler avec des corpus réels de langue, des corpus diversifiés et complémentaires : entendons par là des corpus de langue spécialisée (technique et scientifique), de langue générale, de langue littéraire et de langue orale, puisqu'ils permettent de relever les sens et les sous-sens en usage. Retenons cependant qu'aucun n'est complet en lui-même.

Bien que l'espace ne nous permette pas de traiter en détail des autres informations essentielles à l'article dictionnairique, mentionnons rapidement que notre étude du mot *coulée* a permis de répertorier pas moins de quatorze domaines de spécialité différents. Ceci

un dictionnaire usuel. En ce qui concerne les cooccurrents, nous en avons relevé plus d'une centaine, dont plusieurs sont exclusifs aux corpus québécois. Il en va de même des citations et des exemples recueillis, dont le contenu est propre au contexte québécois. Ces divers éléments lexicaux doivent être considérés en priorité compte tenu de leur adéquation avec le public visé. Il s'agit là d'une riche documentation dont il faudra extraire l'essentiel, à partir de critères préétablis tels que la fréquence, la dispersion dans les corpus, les particularités québécoises, etc. On devra faire de même dans le traitement des mots en concurrence.

D'autres études de mots effectuées dans les mêmes textes et avec la même méthodologie confirment l'importance de prendre en compte des corpus vastes et diversifiés. Par exemple, le groupe complexe *semaine de relâche* ne se trouve que dans les textes de la BDTS; pour sa part, le substantif « contaminant » n'est attesté que dans notre corpus de textes techniques québécois; il est absent des ouvrages tant québécois que français. Or ces mots et expressions sont couramment utilisés au Québec. Enfin, la constitution de ce corpus représentatif de textes techniques québécois nous fournit une première nomenclature des termes techniques généraux devant faire l'objet d'une description dans un dictionnaire général du français québécois. Par ailleurs, les données recueillies sur le mot *coulée* nous permettent d'amorcer une réflexion sur la mise au point d'une méthodologie appropriée en vue du traitement lexicographique de cette portion importante de la langue québécoise.

BOULANGER, Jean-Claude. 1989. « Le traitement des terminologies dans les dictionnaires de langue », *Le Français dans le monde*, 16 p.

CAJOLET-LAGANIÈRE, Hélène et Normand MAILLET. 1995. « Caractérisation des textes techniques québécois », *Présence francophone*, n° 47, p. 113-137.

GAMBIER, Yves. 1995. « Le français dans les communications spécialisées : bilan mitigé », *Présence francophone*, n° 47, p. 10.

**TABLEAU II** **Sens et sous-sens relevés**

| Corpus de textes techniques québécois | Banque de données textuelles de Sherbrooke (BDTS) |
|---|---|
| (1) Le fait de couler, de s'écouler; ce qui s'écoule. | (1) Le fait de couler, de s'écouler; ce qui s'écoule. |
| (2) Action de couler un métal; masse de métal que l'on coule. | (2) Action de couler un métal; masse de métal que l'on coule. |
| (3) **Ravin.** | (3) **Ravin.** |
| (4) **Terrain pâteux répandu en discordance sur d'autres terrains et solidifié par la suite.** | (4) **Terrain pâteux répandu en discordance sur d'autres terrains et solidifié par la suite.** |
| (11) | (11) **Écoulement de la sève d'érable à la suite de l'entaille au printemps.** |
| (12) | (12) **Quantité de sève d'érable qui s'écoule de l'entaille pendant une période de dégel.** |

## TABLEAU I

## Sens et sous-sens relevés

| Corpus de textes techniques québécois | Banque de terminologie du Québec |
|---|---|
| Note 1 : Nous avons mis en gras les sens québécois.<br>Note 2 : Nous conservons la même numérotation des sens dans les cinq tableaux. | |
| (1) Le fait de couler, de s'écouler, ce qui s'écoule. | (1) Le fait de couler, de s'écouler, ce qui s'écoule. |
| (2) Action de couler un métal, masse de métal que l'on coule. | (2) Procédé selon lequel un liquide ou une matière visqueuse est versé, ou introduit d'autre manière, dans un moule ou sur une surface préparée, pour s'y solidifier sans utilisation de pression extérieure. Opération d'évacuation de l'acier liquide du four de fusion par l'ouverture du trou de coulée ménagé à la partie inférieure de ce four et qui passe alors dans la poche en coulant dans un même chenal parmi des briques réfractaires.<br>Action de verser un métal en fusion dans un moule ou une lingotière où il doit se solidifier. Terme de fonderie par lequel on désigne la technique de remplissage de l'empreinte du moule, technique à partir de laquelle est déterminé le tracé des jets de coulée (dans le sens générique du mot). Transvasement de métal liquide d'un four dans une poche, d'une poche dans une autre ou d'une poche dans un moule. Action de faire couler un métal en fusion dans un moule de revêtement par pression mécanique ou force centrifuge. Action de verser dans un moule une matière en fusion (acier, verre, etc.). Action de faire pénétrer un liquide dans un moule. |
| (3) **Ravin** | (3) **Ravin (Dépression allongé d'un terrain).** |
| (4) **Terrain pâteux répandu en discordance sur d'autres terrains et solidifié par la suite.** | (4) |
| (5) | (5) Voie débarrassée d'obstacles, tels que buissons et souches, où passent les chargements halés ou transportés dans une bande de téléphérage. |
| (6) | (6) Corridor dégagé entre deux herbiers aquatiques. |
| (7) | (7) État du métal coulé refroidi. |
| (8) | (8) **Action de faire passer le sirop de l'évaporateur au sirotier.** |
| (9) | (9) **Masse de sirop qui est tirée d'un seul trait de l'évaporateur après avoir atteint le degré de cuisson requis.** |
| (10) | (10) **Période pendant laquelle la sève d'érable s'écoule de l'entaille.** |

**TABLEAU III** **Sens et sous-sens relevés**

| Corpus technique québécois + BDTS | Dictionnaires québécois |
|---|---|
| (1) Le fait de couler, de s'écouler; ce qui s'écoule. | (1) Action de s'écouler, son résultat. (BÉL) |
| (2) Action de couler un métal; masse de métal que l'on coule. | (2) Action de couler un métal; masse de métal que l'on coule. (DFP) (BÉL)<br>Masse de matière en fusion que l'on verse dans un moule. (DQA) |
| (3) **Ravin.** | (3) **Ravin**. (DQA) (BÉL) (GLOSS) |
| (4) **Terrain pâteux répandu en discordance sur d'autres terrains et solidifié par la suite.** | (4) **Terrain pâteux répandu en discordance sur d'autres terrains et solidifié par la suite**. (DFP) |
| (11) **Écoulement de la sève d'érable à la suite de l'entaillage au printemps.** | (11) **Écoulement de la sève d'érable à la suite de l'entaillage au printemps**. (BÉL) (GLOSS) |
| (12) **Quantité de sève d'érable qui s'écoule de l'entaille pendant une période de dégel.** | (12) **Quantité de sève d'érable qui s'écoule de l'entaille pendant un temps déterminé**. (GLOSS) |
| (13) | (13) **Petit cours d'eau qui coule au fond d'un ravin**. (DQA) |
| (14) | (14) Trace laissée par le passage d'un animal dans des buissons, un sous-bois, etc. (DFP) (BÉL)<br>Chemin étroit que le cerf suit pour se rendre dans son réduit. (BÉL) |
| (15) | (15) Sorte d'écriture penchée dont toutes les lettres sont unies par des liaisons. (BÉL) |

## TABLEAU IV

### Sens et sous-sens relevés

| **Corpus technique québécois + BDTS + dict. québécois** | **Dictionnaires français usuels** *(Petit Robert, Petit Larousse illustré)* |
|---|---|
| (1) Le fait de couler, de s'écouler; ce qui s'écoule. | (1) Le fait de couler, de s'écouler; ce qui s'écoule. |
| (2) Action de couler un métal; masse de métal que l'on coule. | (2) Action de verser du métal en fusion dans un moule; masse de métal ainsi versée. |
| (3) **Ravin.** | (3) |
| (4) **Terrain pâteux répandu en discordance sur d'autres terrains et solidifié par la suite.** | (4) |
| (11) **Écoulement de la sève d'érable à la suite de l'entaillage au printemps.** | (11) |
| (12) **Quantité de sève d'érable qui s'écoule de l'entaille pendant une période de dégel.** | (12) |
| (13) **Petit cours d'eau qui coule au fond d'un ravin.** | (13) |
| (14) Trace laissée par le passage d'un animal dans des buissons, un sous-bois, etc. | (14) Sentier étroit par lequel le cerf gagne son réduit. |
| (15) Sorte d'écriture penchée dont toutes les lettres sont unies par des liaisons. | (15) |
| (16) Masse de matière en fusion que l'on verse dans un moule. | (16) Masse de matière en fusion que l'on verse dans un moule. Masse de matière plus ou moins liquide ou en fusion qui s'écoule, se répand. |
| (17) | (17) Petit chemin. |
| (18) | (18) Ton qui forme une tache dominante dans un tableau. |
| (19) | (19) Déplacement méridien d'une masse d'air perpendiculairement aux flux zonaux de la circulation atmosphérique générale. |
| (20) | (20) Action de verser du verre en fusion sur une table en fonte. |

## TABLEAU V

## Sens et sous-sens relevés

| Corpus technique + BDTS + dict. québécois + franc. | Grands dictionnaires français *(GLLF, TLF)* |
|---|---|
| (1) Le fait de couler, de s'écouler; ce qui s'écoule. | (1) Le fait de couler, de s'écouler, ce qui s'écoule. |
| (2) Action de verser du métal en fusion dans un moule; masse de métal ainsi versée. | (2) Action de verser du métal en fusion dans un moule, masse de métal ainsi versée. Dans le clichage typographique, action de verser du métal en fusion sur l'empreinte placée dans le moule. |
| (3) **Ravin.** | (3) |
| (4) **Terrain pâteux répandu en discordance sur d'autres terrains et solidifié par la suite.** | (4) |
| (11) **Écoulement de la sève d'érable à la suite de l'entaillage au printemps.** | (11) |
| (12) **Quantité de sève d'érable qui s'écoule de l'entaille pendant une période de dégel.** | (12) |
| (13) **Petit cours d'eau qui coule au fond d'un ravin.** | (13) |
| (14) Trace laissée par le passage d'un animal dans des buissons, un sous-bois, etc. Sentier étroit par lequel le cerf gagne son réduit. | (14) Passage étroit dans la forêt suivi ou tracé par l'animal. |
| (15) Sorte d'écriture penchée dont toutes les lettres sont unies par des liaisons. | (15) Écriture penchée de droite à gauche dont les lettres aux jambages parallèles sont liées entre elles. |
| (16) Masse de matière en fusion que l'on verse dans un moule. Masse de matière plus ou moins liquide ou en fusion qui s'écoule, se répand. | (16) Masse de matière en fusion que l'on verse dans un moule. |
| (17) Petit chemin. | (17) |
| (18) Ton qui forme une tache dominante dans un tableau. | (18) |
| (19) Déplacement méridien d'une masse d'air perpendiculairement aux flux zonaux de la circulation atmosphérique générale. | (19) Déplacement perpendiculairement aux flux zonaux de la circulation atmosphérique générale. |
| (20) Action de verser du verre en fusion sur une table en fonte. | (20) |
| | (21) Par analogie, se dit d'une foule en mouvement. |
| | (22) Par analogie, en parlant du temps. |
| | (23) Orifice en forme de conduit, permettant au métal fondu d'atteindre l'empreinte même du moule. |
| | (24) Greffe en coulée. |
| | (25) Courbure des fonds extérieurs de la carène d'un navire. |
| | (26) Glissée immergée du nageur en position allongée. |
| | (27) Laps de temps pendant lequel une ligne flottante dérive. |
| | (28) Espace parcouru par la ligne flottante dérivant dans un courant. |

Julie Auger et Yvan Rose (réd.). (1997). *Explorations du lexique*. Québec: CIRAL, 281-292.

# Quelques critères de sélection des termes dans les dictionnaires bilingues

*Aline Francoeur et René Gemme*
Université de Montréal

## Introduction

Depuis des siècles, les dictionnaires de langue, tant bilingues qu'unilingues, recensent des termes techniques et scientifiques[1] dans leur nomenclature. Comme il est cependant impossible de consigner dans ces ouvrages l'ensemble des termes techniques et scientifiques existants, le lexicographe se doit de procéder à une sélection rigoureuse et souvent difficile. Dans le cadre de notre étude, nous nous sommes intéressés à la sélection des termes techniques et scientifiques dans les dictionnaires de langue avec deux objectifs bien précis : a) dégager les critères qui guident le lexicographe dans ses décisions quant à l'inclusion ou au rejet des termes; b) identifier les critères de sélection les plus importants à prendre en compte dans le cadre d'un projet dictionnairique bien précis, à savoir le projet de *Lexicographie comparée du français et de l'anglais*, que nous décrirons brièvement plus loin.

## 1. Généralités sur la sélection des termes

La question de la sélection des termes dans les dictionnaires de langue n'a pas fait l'objet d'études systématiques et les rares auteurs qui s'y sont intéressés reconnaissent qu'elle soulève un certain nombre de problèmes, mais proposent rarement des solutions. En fait, on constate que l'objectivité des lexicographes quant au choix des termes est souvent mise en doute. Autrement dit, on soupçonne le lexicographe d'agir de façon arbitraire, souvent involontairement, parce qu'il

---

[1] La définition du terme technique et scientifique que nous adoptons dans le cadre de notre étude a été proposée par Louis Guilbert en 1975 pour ce qu'il désignait alors sous le nom de *terminologisme* (Boulanger 1986:63). En vertu de cette définition, le terme technique et scientifique est une *unité lexicale appartenant aux terminologies spécialisées des sciences et des techniques*.

est en quelque sorte limité par ses connaissances, sa culture, son milieu ou son éducation, mais parfois aussi volontairement, parce que le terme appartient à un domaine pour lequel il éprouve un grand intérêt (Béjoint 1988:361). C'est donc dire que les lexicographes ne disposent pas, de façon générale, de critères solides sur lesquels appuyer leurs décisions.

**2. Discussion de quelques critères de sélection des termes**

Les critères de sélection que nous allons présenter ont été élaborés dans l'esprit d'un projet dictionnairique bien précis, à savoir le projet de *Lexicographie comparée du français et de l'anglais*, mieux connu sous le nom de *Dictionnaire bilingue canadien* (DBC). Ce projet, mené conjointement à l'Université d'Ottawa, l'Université de Montréal et l'Université Laval, a pour objectif premier la description systématique de l'anglais et du français en usage au Canada[2]. En vertu du programme dictionnairique adopté, le DBC sera un dictionnaire descriptif bidirectionnel d'encodage et de décodage qui devrait compter approximativement 100 000 entrées. Il s'adressera principalement à un public de francophones et d'anglophones qui possèdent déjà une bonne connaissance de leur langue seconde (traducteurs, journalistes, universitaires, etc.).

**2.1 Diffusion sociale du terme**

Nous croyons qu'il serait important de retenir dans la nomenclature du DBC les termes qui présentent un certain degré de diffusion auprès du public-cible du dictionnaire. Par CERTAIN DEGRÉ DE DIFFUSION, nous entendons que le terme, et donc la notion à laquelle il renvoie, est assez bien connu de l'utilisateur du dictionnaire. On peut penser à des termes comme *sida*, *logiciel*, *cardamome*, etc., qui relèvent de domaines d'activités de la vie quotidienne ou qui sont largement véhiculés par les médias. Pour mesurer la diffusion sociale

[2] Jusqu'à maintenant, cela n'a été fait que partiellement dans un ouvrage aujourd'hui épuisé, le *Dictionnaire canadien*, produit sous la direction de Jean-Paul Vinay et publié en 1962.

des termes, les dictionnairistes s'appuient principalement sur l'utilisation de bases de données textuelles informatisées (BDTI). Ces dernières, connues aussi sous le nom de corpus informatisé, interviennent à plusieurs niveaux dans l'élaboration des dictionnaires[3], notamment dans l'établissement des nomenclatures, et c'est tout particulièrement cet aspect qui nous intéresse. L'équipe du DBC dispose depuis 1994 d'un corpus informatisé baptisé TEXTUM[4] qui contient à l'heure actuelle quelque 315 millions d'unités lexicales, soit 105 millions d'unités lexicales françaises, et 210 millions d'unités lexicales anglaises, réparties comme suit :

**Tableau 1**. Répartition des sources dans la base de données textuelles informatisée TEXTUM

| | | |
|---|---|---|
| textes journalistiques | anglais | Canada : *The Gazette, The Ottawa Citizen, The Toronto Star, The Daily News* (Halifax), *Calgary Herald* et *The Vancouver Sun*<br>Etats-Unis : *The Wall Street Journal* |
| | français | Canada : *La Presse, Le Soleil, Le Droit, l'Actualité, Le Devoir*<br>France : *Le Monde, Ouest France* |
| textes littéraires | anglais | Canada : romans d'auteurs canadiens-anglais fournis par l'Université Queens |
| | français | Canada : romans publiés par la maison d'édition Leméac |

[3] Les BDTI permettent notamment d'attester les emplois du mot traité, de trouver des exemples d'emplois particuliers souvent non répertoriés dans les dictionnaires existants, de relever les constructions syntaxiques, les collocations et les expressions dans lesquelles le mot traité apparaît le plus souvent.

[4] TEXTUM, développé par Jean Baudot en collaboration avec les Services informatiques de l'Université de Montréal, utilise le système d'interrogation PAT, créé à l'Université de Waterloo (Ontario) pour l'informatisation du *Oxford English Dictionary*.

| | | |
|---|---|---|
| textes scientifiques et techniques | anglais | Canada : revue *Canadian Geographic*<br>États-Unis : textes gouvernementaux provenant du Department of Energy |
| | français | Canada : revue *Interface*, publiée par l'Association canadienne-française pour l'avancement des sciences |

Il va sans dire que pour tenir compte de la diffusion sociale du terme, il faudra au préalable définir des règles qui permettront de fixer, en quelque sorte, ce qui sera considéré comme un CERTAIN DEGRÉ DE DIFFUSION. On devra déterminer un nombre précis de sources dans lesquelles le terme devra apparaître et délimiter également un nombre minimal d'occurrences du terme dans chacune des sources. Il est important de souligner que la diversité des sources est plus importante à considérer que le nombre d'occurrences du terme qui fait l'objet de la demande. De plus, la nature des sources devrait être prise en compte pour décider du sort du terme. C'est pourquoi nous considérons qu'il serait important que les termes techniques et scientifiques retenus dans le DBC figurent dans au moins deux sources différentes de TEXTUM et, plus important encore, que l'une de ces sources soit canadienne.

**2.2 Domaine d'appartenance du terme**

En vertu du programme dictionnairique adopté, le DBC sera adapté à la communauté canadienne en particulier, et, de façon plus générale, à la communauté nord-américaine. Par conséquent, on tiendra compte du contexte social, économique, politique et culturel dans lequel vivent les Canadiens et les Nord-Américains en général. C'est donc dire qu'on accordera une attention particulière aux termes dénommant les sports qu'ils pratiquent, les conditions climatiques dans lesquelles ils vivent, la faune et la flore qui les entoure, etc. On trouvera donc dans le DBC des termes comme *high sticking* (domaine du hockey), *brushwolf*, terme canadien désignant le *coyote*, *sanguinaire*, nom d'une plante qui pousse dans la région du Bas-Saint-Laurent, etc. De même, on délimitera certains domaines considérés

comme étant d'intérêt particulier pour les Canadiens; on peut penser, par exemple, à l'industrie de la pêche, à l'hydroélectricité, à l'exploitation minière, etc., domaines d'activités majeurs en contexte canadien. Il faut cependant insister sur l'importance de dresser au préalable une liste exhaustive des domaines qui seront considérés comme étant d'intérêt particulier, cela, bien sûr, afin d'éviter toute subjectivité dans le choix des domaines en question. Cette liste pourrait aussi intervenir dans la sélection du contenu de la base de données textuelles développée dans le cadre du projet. Par ailleurs, nous ne croyons pas nécessaire de retenir l'ensemble des termes relevant des domaines d'intérêt particulier préalablement identifiés compte tenu qu'un certain nombre d'entre eux seront sans doute d'un degré de technicité ou de scientificité trop élevé pour apparaître dans un dictionnaire de langue. On peut penser, dans le domaine de l'hydroélectricité, au terme *barrage à contreforts à voûtes multiples*, terme hautement technique par rapport aux termes *barrage à contreforts* et *barrages à voûtes*. Par conséquent, nous croyons qu'il faudrait prioritairement prendre en compte la diffusion sociale et retenir seulement les termes les plus répandus, donc, ceux qui présentent un certain nombre d'occurrences dans TEXTUM.

**2.3 Consultation de dictionnaires existants**

La nomenclature de base du DBC a été établie à partir de quatre dictionnaires unilingues élaborés au Canada : le *Gage Canadian Dictionary* et le *Penguin Canadian Dictionary* pour la nomenclature anglaise, le *Dictionnaire du français Plus* et le *Dictionnaire québécois d'aujourd'hui* pour la nomenclature française (Baudot *et al.* 1993:11). Bien entendu, la liste de mots dressée à partir de ces ouvrages sera adaptée; certaines catégories de mots seront systématiquement exclues du DBC, d'autres seront ajoutées.[5] Nous serons de plus amenés à

[5] Les mots désuets, les variantes régionales autres que canadiennes et les termes hautement techniques et scientifiques seront exclus systématiquement. Par ailleurs, on ajoutera des mots de la langue générale absents des dictionnaires de base consultés mais présentant un certain nombre d'occurrences dans TEXTUM, des

consulter les dictionnaires bilingues anglais-français / français-anglais existants en partant du principe que pour faire mieux qu'eux, il faut au moins faire autant. Par contre, l'orientation de certains de ces ouvrages étant trop différente de celle du DBC, nous ne croyons pas qu'il soit utile de les prendre tous en considération. À notre avis, deux ouvrages suffiraient pour établir la nomenclature technique et scientifique de départ du DBC : le *Dictionnaire Hachette-Oxford français-anglais / anglais-français* et *Le Robert & Collins Super Senior*. Par contre, ce ne sont pas tous les termes techniques et scientifiques recensés dans ces ouvrages qui devraient être retenus dans le DBC; nous ne voyons pas l'intérêt, en effet, de recenser des termes qui ne sont pas en usage au Canada ou dont le degré de diffusion laisse croire qu'ils sont peu connus du public-cible. C'est donc dire que, encore une fois, le corpus devrait trancher quant au sort réservé au terme.

**2.4 Uniformité dans la nomenclature**

Comme on le sait, il est impossible de recenser l'ensemble des noms de genres ou d'espèces dans un dictionnaire de langue, quel qu'il soit. Henri Cottez (1994:22) estime cependant qu'il est essentiel de retenir les termes qui désignent des regroupements en classes. Selon cette approche, il serait possible d'obtenir une certaine uniformité dans les nomenclatures de dictionnaires. Il existe un nombre important de termes qui désignent des regroupements en classes. Prenons par exemple quelques-uns des termes français qui désignent les grandes familles d'ongulés (mammifères pourvus de sabots), à savoir les termes *cervidé, giraffiné, bovinè, antilopiné, capriné, équidé, camélidé, chevrotain*. Si l'on suit le raisonnement d'Henri Cottez, chacun de ces termes devrait être retenu puisqu'il désigne un regroupement en

---

canadianismes proposés par les spécialistes du français et de l'anglais canadiens qui participent au projet, des acronymes et noms propres représentatifs de réalités canadiennes et, finalement, des termes techniques appartenant à des domaines d'intérêt particulier pour les Canadiens, comme nous l'avons déjà mentionné plus haut.

classes, ces classes se subdivisant ensuite en diverses espèces. De même, les termes anglais qui désignent les grandes classes de médicaments qui agissent sur le système nerveux central, à savoir les termes *analgesic*, *anticonvulsant*, *antidepressant*, *antiparkinsonism*, *antipsychotic*, *anxiolytic* et *sedative* devraient tous être retenus puisqu'ils renvoient à des grands regroupements dans le domaine de la pharmacologie.

À notre avis, le premier problème qui se pose avec l'approche d'Henri Cottez est de savoir jusqu'où aller dans la description des termes qui désignent des grandes classes d'objets. Il faudrait, toujours selon cette approche, que tous les cohyponymes d'un même hyperonyme qui désignent un grand regroupement d'objets apparaissent en nomenclature du dictionnaire. Ainsi, la famille des ongulés se subdivisant en 17 grandes classes, chacun des termes qui désignent ces classes devrait apparaître dans le dictionnaire. Mais si ensuite chacune de ces classes se subdivise à son tour en classes plus spécifiques, doit-on aussi traiter les termes qui désignent ces dernières? Où s'arrête la description? On peut de plus se demander s'il est vraiment pertinent de recenser dans un dictionnaire de langue des termes que seul un nombre restreint de spécialistes connaissent, c'est-à-dire des termes qui présentent une faible diffusion auprès de l'utilisateur-cible du dictionnaire. Nous avons d'ailleurs tenté une expérience afin de pouvoir mieux juger de toute cette question. Nous avons interrogé, dans quatre sources anglaises de TEXTUM, les termes désignant les classes de médicaments agissant sur le système nerveux central. Nous avons également interrogé quatre sources françaises de TEXTUM afin de voir si certains termes désignant des classes d'ongulés y étaient représentés. Les résultats de nos recherches[6] sont consignés dans les deux tableaux suivants :

---

[6]Pour ce qui est des termes anglais désignant les médicaments qui agissent sur le système nerveux central, on trouve un grand nombre d'occurrences en emploi adjectival, du type *sedative effect*, *antidepressant effect*, etc. Ces occurrences n'ont pas été comptabilisées, le terme ne désignant pas alors la classe de médicaments.

**Tableau 2** Nombre d'occurrences des termes qui renvoient aux classes de médicaments agissant sur le système nerveux central dans quatre sources anglaises de TEXTUM

| | Queens | Wall Street Journal | Department of Energy | English Canada Press |
|---|---|---|---|---|
| analgesic | 9 | 13 | 4 | 29 |
| anticonvulsant | 0 | 0 | 6 | 2 |
| antidepressant | 0 | 3 | 12 | 39 |
| antiparkinsonism | 0 | 0 | 0 | 0 |
| antipsychotic | 0 | 0 | 2 | 0 |
| anxiolytic | 0 | 1 | 2 | 0 |
| sedative | 2 | 13 | 2 | 91 |

**Tableau 3** Nombre d'occurrences de certains termes qui renvoient aux grandes classes d'ongulés dans quatre sources françaises de TEXTUM

| | Presse canadienne française | Ouest France | Interface | Le Monde |
|---|---|---|---|---|
| cervidé | 119 | 4 | 0 | 4 |
| giraffiné | 0 | 0 | 0 | 0 |
| bovine | 1 | 0 | 0 | 0 |
| antilopiné | 0 | 0 | 0 | 0 |
| équidé | 3 | 2 | 0 | 2 |
| camélidé | 3 | 1 | 0 | 2 |
| chevrotain | 0 | 0 | 0 | 0 |
| caprine | 0 | 0 | 0 | 0 |

Bien entendu, nous ne pouvons nous permettre de tirer de

conclusions à partir de ces deux seuls exemples. Par contre, à la lumière des résultats obtenus, nous constatons que parmi les termes qui ont fait l'objet d'une interrogation dans TEXTUM, certains semblent définitivement plus répandus que d'autres dans la langue générale, c'est le cas notamment des termes *analgesic*, *antidepressant* et *sedative* en ce qui concerne la première série de termes, et des termes *cervidé*, *équidé* et *camélidé* pour ce qui est des termes désignant les familles d'ongulés.

À première vue, l'uniformité peut sembler un critère de sélection fort intéressant à prendre en compte. Nous croyons cependant qu'il est difficile d'envisager concrètement l'application d'une telle approche compte tenu, notamment, des contraintes de temps, et surtout, des contraintes d'espace auxquelles les lexicographes doivent généralement se soumettre. Par conséquent, c'est encore une fois la diffusion sociale du terme qui, à notre avis, devrait être considérée d'abord et avant tout.

**2.5 Concept dénommé par le terme**

Selon Henri Cottez (1994:20), il faudrait retenir tout terme technique ou scienfique qui renvoie à une nouvelle réalité susceptible de changer notre vision du monde. Suivant ce principe, nous croyons qu'il serait effectivement important de recenser tout néologisme qui désigne une nouvelle technologie, une nouvelle science, une nouvelle activité spécialisée pouvant influencer notre vision du monde ou changer notre existence quotidienne. On peut penser, par exemple, aux nombreux néologismes qui sont apparus récemment dans le domaine de l'informatique avec l'avènement de l'autoroute électronique (*internaute*, *logiciel de navigation*, *Netiquette*, etc.). Ces termes désignent des notions qui sont effectivement susceptibles de changer le quotidien d'un nombre relativement important de consommateurs, et qui d'ailleurs s'implantent à un rythme extrêmement rapide. De tels néologismes devraient à notre avis être retenus.

**2.6 Officialisation du terme par un organisme reconnu**

Il pourrait, dans certains cas, être intéressant de tenir compte

de l'officialisation du terme dans le processus d'établissement des nomenclatures scientifiques et techniques. On peut penser à un cas bien précis, à savoir lorsque le lexicographe est confronté à une série synonymique et ne peut retenir chacun des synonymes. Par conséquent, il pourrait décider de retenir, le cas échéant, le terme qui aurait fait l'objet d'une officialisation par un organisme compétent (OLF, AFNOR, etc.). En adoptant ce critère, le lexicographe ferait en sorte que le dictionnaire serve d'outil de diffusion et d'implantation des termes officiels.

**Conclusion**

Notre étude a permis de constater que, de façon générale, la sélection des termes techniques et scientifiques pour l'élaboration des nomenclatures de dictionnaires de langue se fait souvent de façon arbitraire et aléatoire. Compte tenu de cette situation, nous croyons qu'il serait important de définir des critères précis pouvant guider les lexicographes dans leur travail. Il ne nous semble toutefois pas possible d'élaborer une grille de critères généraux qui soient applicables indifféremment à tout projet dictionnairique; à notre avis, chaque équipe lexicographique devrait définir ses propres critères de sélection des termes techniques et scientifiques en fonction du programme dictionnairique qu'elle aura adopté.

Dans cet ordre d'idées, nous avons identifié un certain nombre de critères qui devraient selon nous être pris en considération dans le cadre du projet du *Dictionnaire bilingue canadien*. D'abord, nous croyons que la nomenclature technique et scientifique de base du DBC pourrait être élaborée à partir des nomenclatures de deux dictionnaires bilingues existants, à savoir le *Dictionnaire Hachette-Oxford* et le *Robert & Collins Super Senior*. Par contre, il faudra adapter cette nomenclature aux buts spécifiques du DBC. Pour ce faire, il faudra notamment tenir compte de la diffusion sociale du terme en se basant sur les résultats d'interrogations de la base de données textuelles informatisée TEXTUM. De plus, il nous paraît important d'accorder une attention particulière aux termes relevant de domaines d'intérêt particulier pour la population canadienne, cela sous-entend qu'il faudra

au préalable dresser une liste exhaustive des domaines que l'on jugera d'intérêt pour le futur utilisateur du DBC. De même, nous pensons qu'il faudrait prendre en compte le concept auquel le terme renvoie et retenir tout néologisme désignant une nouvelle technologie, une nouvelle science, une nouvelle activité susceptible d'influencer notre vision du monde et de changer nos habitudes de vie. En outre, dans certains cas particuliers, il pourra s'avérer utile de tenir compte de l'officialisation du terme par un organisme reconnu.

En terminant, bien que nous ne nous soyons pas penchés plus à fond sur la question, nous croyons que les critères que nous avons isolés pourraient aussi s'appliquer au processus de sélection des termes qui seront traités en sous-entrées dans le DBC.

## Références

### 1) Monographies et articles

Baudot, Jean, Lionel Boisvert, Jacqueline Bossé-Andrieu, Jean-Claude Boulanger, André Clas, Monique C. Cormier & Roda P. Roberts. 1993. *Dictionnaire bilingue canadien : description du projet*.

Béjoint, Henri. 1988. "Scientific and Technical Words in General Dictionaries". *International Journal of Lexicography*1, 4: 354-368.

Boulanger, Jean-Claude. 1986. *Aspects de l'interdiction dans la lexicographie française contemporaine*. Tubingen: Max Niemeyer Verlag.

Cottez, Henri. 1994. «La place des termes techniques et scientifiques dans un dictionnaire de langue». Danielle Candel (dir.). *Français scientifique et technique et dictionnaire de langue*. Paris: Didier Érudition, pp. 13-26.

Magai, Tamás. 1983. "Technical and General: Problems of Vocabulary Selection in a Medium-Size Bilingual Dictionary". In *LEXeter 1983 Proceedings*. Tübingen: Max Niemeyer Verlag, pp. 221-225.

**2) Dictionnaires bilingues**

*Dictionnaire Hachette-Oxford français-anglais / anglais-français*. 1994. Oxford: Oxford University Press.

*Grand Dictionnaire français-anglais / anglais-français*. 1993. Paris: Larousse.

*Harrap's Shorter French and English Dictionary*. 1993. New York: Harrap BooksLimited.

*Le Robert & Collins Super Senior*. 1995. Paris: HarperCollins Publishers et Dictionnaires Le Robert.

**3) Dictionnaires unilingues anglais**

*Collins English Dictionary*. 1991. Glasgow: Harper Collins Publishers.

*Gage Canadian Dictionary*.1983. Toronto: Gage Publishing Company.

*The Penguin Canadian Dictionary*. 1990. Penguin Books Canada / Copp Clark Pitman Book.

**4) Dictionnaires unilingues français**

*Dictionnaire du français Plus*. 1988. Montréal: CEC.

*Dictionnaire québécois d'aujourd'hui*. 1993. Saint-Laurent: Dicorobert.

*Le Nouveau Petit Robert*. 1993. Paris: Le Robert.

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 293-304.

# Inventaire du vocabulaire de l'alimentation

*Josée Mercier*
Université de Sherbrooke

## INTRODUCTION

Les problèmes qui émanent de ce que *Gattel*[1] proposa d'appeler *nomenclature* ont toujours préoccupé les lexicographes. Le repérage des unités lexicales dans une recherche lexicographique constitue une démarche cruciale, pour le moins complexe. Elle regroupe deux étapes essentielles : la cueillette et la sélection des unités.

L'absence de méthode uniforme, entre les divers membres d'une recherche lexicographique, a souvent été dénoncée comme étant l'une des causes majeures qui perturbent l'équilibre, la cohérence de la nomenclature d'un dictionnaire. Outre l'absence de méthode, plusieurs critères peuvent faire varier les aspects quantitatif et qualitatif d'une nomenclature. À titre d'exemples, mentionnons l'organisation de la nomenclature, les modalités de présentation, les mots oubliés, la valeur des exemples, les sens incomplets ou absents, etc.

Ces quelques points prouvent à quel point il est important d'établir une méthode d'inventaire qui garantisse aux chercheurs une approche plus objective et plus systématique de la masse lexicale d'un domaine.

---

1

« *Dic. Universel* (1813), Préf., xj. Le mot est entré dans la dernière édition du TRÉVOUX (1771) avec cette valeur nouvelle. On lit dans la *Préf.*, p. vii : " on a d'abord consulté tous les autres dictionnaires, pour qu'il n'échappât, s'il étoit possible, aucun des mots de la Langue, et pour former la nomenclature la plus riche et la plus étendue " » (QUEMADA 1967 : 265)

Pour entamer cette étude, les deux dictionnaires québécois, le *Dictionnaire québécois d'aujourd'hui* 1993 (DQA) et le *Dictionnaire du français plus à l'usage des francophones d'Amérique* 1988 (DFP), ont servi de point de départ dans l'établissement de notre problématique et dans l'enrichissement de notre réflexion. Par ailleurs, les remarques qui seront données à partir de ces deux dictionnaires ne visent pas à critiquer négativement le travail qui a été fait dans ces deux ouvrages, au contraire.

## 1.1 LIMITES DES DICTIONNAIRES : ASPECTS QUANTITATIFS[2]

### 1.1.1 Mots oubliés

« Les mots oubliés sont ceux que l'auteur n'a ni présentés ni définis dans la liste des adresses mais qui apparaissaient dans le contenu des articles » (QUEMADA 1967 : 285). Habituellement, tous les mots présents dans les définitions doivent se retrouver définis, à leur tour, dans la nomenclature.

Malgré tout, nous avons constaté certaines *infractions* à cette règle qui entravent un principe fondamental en discours : « ne rien laisser qui ne soit expliqué et défini » (DUBOIS Jean et Claude 1971 : 58). Par exemple, dans le *DQA*, sous le vocable *caroubier*, nous lisons : « Arbre à feuilles persistantes, à fleurs rougeâtres, qui produit un fruit sucré (appelé *caroube* ou *carouge*, n.f.) ». Or *caroube* ou *carouge* sont absents de la nomenclature.

### 1.1.2 Mots absents, mots douteux

La nomenclature d'un dictionnaire doit répondre aux besoins des utilisateurs. Or dans les dictionnaires québécois actuels (*DQA* et *DFP*), certains vocables, du domaine de l'alimentation, ne sont peut-

---

2 Bernard Quemada a étudié les aspects quantitatifs et qualitatifs de la nomenclature de différents dictionnaires (QUEMADA 1967).

être pas représentatifs des usages linguistiques québécois. Parmi ces mots, nous retrouvons : *achar* (*achard*), *alkermès*, *garbure*, *halva*, *panatela* (*panatella*), etc. D'autres vocables, par contre, sont absents du *DQA* et du *DFP*, mais sont bien attestés au Québec : *brownies*, *chaudrée*, *chutney*, *chiche-kebab*, *coupe glacée*, *pâté impérial*, *strudel*, etc. Enfin, certains mets, d'origines diverses, se retrouvent sur nos tables et sont présents dans le *DQA* et le *DFP* : le *baklava*, la *bavette*, la *choucroute*, le *gaspacho*, le *magret de canard*, etc. en sont quelques exemples.

1.1.3 Mots ou sens redoublés

Nous nous sommes aperçue, qu'à l'intérieur d'une même nomenclature, certains groupes complexes se retrouvent en sous-entrées à deux endroits à la fois, faisant varier légèrement l'aspect quantitatif de la nomenclature. Par exemple, sous l'entrée *filet* du *DQA*, nous lisons : « *Un filet mignon*, la pointe du filet de boeuf. » et, toujours dans le même dictionnaire, sous *mignon* : « FILET MIGNON : bifteck coupé dans la pointe du filet. »

1.1.4 Nombre d'entrées et richesse

La qualité d'un dictionnaire ne se mesure pas par son nombre d'entrées, c'est bien connu. « Si la notion de complétude n'a aucun sens, celle du nombre de mots traités n'a aucune pertinence. » (NPR : XI). Toutefois, il n'y a pas que cet aspect qui limite la représentativité de la valeur d'un dictionnaire. Le traitement des homographes et des polysèmes, la disposition des mots dérivés, les variantes orthographiques, etc, s'avèrent autant de prétextes qui peuvent augmenter la proportion des entrées dans un dictionnaire. Pour n'en donner qu'un exemple, les modalités de présentation varient suffisament entre le *DQA* et le *DFP* pour que le dénombrement des entrées soit différent : le *DQA* privilégie la polysémie, alors que le *DFP* opte pour l'homonymie, les auteurs du *DQA* ont préféré un regroupement des familles de mots sous l'entrée principale, alors que ceux du *DFP* ont choisi l'approche contraire.

## 1.2 LIMITES DES DICTIONNAIRES : ASPECTS QUALITATIFS

### 1.2.1 Difficultés de repérage

Il est parfois difficile, à la seule lecture de la définition, de déduire si un fruit, une plante, un champignon, etc. s'avère comestible, ce dernier trait nous permet de savoir s'il doit être retenu à la nomenclature de l'alimentation. Certaines **définitions** nous paraissent donc **incomplètes**. En voici quelques exemples tirés du *DFP* : *baguenaude* « fruit du baguenaudier, ayant la forme d'une vessie quadrangulaire », *balise* « fruit du balisier, dont la graine fournie un colorant pourpre », *souchette* « sorte de collybie poussant sur les souches », *sureau* « arbuste (genre *Sambucus*, fam. caprifoliacées) dont les fleurs, hermaphrodites et regroupées en corymbe, donnent un fruit noir ou rouge et dont le bois renferme un large canal médullaire », etc. Après des recherches plus approfondies, notamment dans des ouvrages encyclopédiques, il ressort que seuls la *souchette* et les fruits du *sureau* sont comestibles. Nous avons remarqué que ce phénomène se produisait souvent avec les définitions faites sur le modèle de « fruit de... » et avec les aliments qui sont consommés plus rarement (abats, gibier, algues comestibles, fruits et légumes exotiques ou sauvages, etc.).

### 1.2.2 Valeur des exemples

Il est intéressant de remarquer que la comestibilité est évoquée ailleurs que dans les définitions. Ainsi, le vocable *avoine* provenant du *DQA* mentionne la définition suivante : « Plante graminée (céréale) dont le grain sert surtout à l'alimentation des chevaux et des volailles. *Des céréales d'avoine. Un biscuit au son d'avoine.* » Néanmoins, les exemples donnés par le *DQA* ne représentent-ils pas uniquement l'alimentation humaine?

### 1.2.3 Sens absents

Par ailleurs, d'autres **sens** courants sont **absents**, soit du *DQA*, soit du *DFP*. Par exemple, au mot *trempette* dans le *DFP*, le sens

d'une *préparation de sauce assaisonnée* est manquant et sous le vocable *beignet*, nous ne retrouvons pas le sens d'un *petit beigne*.

## 2.1 HYPOTHÈSE DE TRAVAIL

L'établissement d'une nomenclature de dictionnaire peut se faire de plusieurs façons. Elle peut s'effectuer à partir de d'autres nomenclatures déjà établies (*DQA*, *DFP*), elle peut faire appel à une documentation plus variée (*TLF*), elle peut se réaliser à partir de la compétence et de l'intuition des auteurs (*DFC*, *Dictionnaire Bordas*) ou encore elle peut opter pour une méthode mixte (*Lexis*, *Le Petit Robert*).

À notre avis, élaborer une nouvelle nomenclature à partir de nomenclatures déjà existantes, constitue un bon point de départ. Cette méthode demeure cependant incomplète et devra être enrichie.

## 3.1 OBJECTIFS

L'objet de la présente étude se situe à l'intérieur du projet de description des usages du français québécois, telle que poursuivie par le Centre d'analyse et de traitement informatique du français québécois (CATIFQ), et **consiste en l'établissement d'une nomenclature originale d'une partie du vocabulaire de l'alimentation**. Naturellement, pour y parvenir, nous avons adopté un cadre lexicographique qui soit plus large que celui d'un dictionnaire général. Par exemple, dans cette nomenclature, nous retrouvons des particularités gastronomiques qu'une nomenclature dictionnairique ne conserverait peut-être pas (algues comestibles, aux fruits, légumes ou végétaux exotiques ou sauvages, etc.).

### 3.1.1 Objectif secondaire

À l'intérieur de notre nomenclature, nous voulons dégager les entrées susceptibles d'être représentées dans un dictionnaire québécois

général. Ceci représente une première approche, la liste proposée ne sera donc pas définitive.

## 4.1 MÉTHODOLOGIE

### 4.1.1 Démarche circulaire

En élaborant notre nouvelle nomenclature, nous avons dû franchir plusieurs étapes, les unes étant parfois préalables, parfois complémentaires aux autres. Ces étapes viennent préciser notre démarche de dépouillement et s'insèrent dans un processus circulaire. Elles possèdent un caractère provisoire. Par exemple, nous savons que pour dresser la liste des mots appartenant à l'alimentation, il faut avoir, au départ, une idée du champ de travail à délimiter. Pour se faire, nous devons tracer les grandes lignes de ce dernier : aromates, céréales, confiseries, fruits, légumes, viandes, etc. Toutefois, c'est uniquement en dépouillant les ouvrages que nous pouvons vérifier l'exactitude de notre délimitation et au besoin, la compléter. Il en est de même pour l'élaboration de l'arbre du domaine et bien sûr, pour l'établissement de la nomenclature.

### 4.1.2 Délimitation du champ de travail

Comme nous l'avons mentionné antérieurement, avant de pouvoir élaborer une nomenclature par domaine d'activité, il faut franchir certaines étapes. La délimitation du champ de travail en est une. Avant d'aller plus loin, il importe de préciser que notre domaine d'étude de recherche correspond à **l'alimentation humaine.** Nous avons donc écarté les mots faisant référence à des aliments « religieux » (*hostie, manne*, etc.), mythologiques (*ambroisie, népenthèse*, etc.), magiques (*potion, philtre*, etc.) et, bien sûr, à l'alimentation animale. Ainsi, pour la présente étude, la liste des mots appartenant au domaine de l'alimentation regroupe essentiellement tous les liquides qui se boivent et tous les aliments qui se mangent. Nous avons exclu la faune marine (poissons, crustacés, etc.), d'autres chercheurs travaillant déjà dans ce sous-domaine, la chimie alimentaire, dont la grande technicité nous paraissait inaccessible, et les unités qui ne

représentaient pas des aliments identifiables tels que le contenu des contenant, les quantités de nourriture, les appellations de repas, les familles de végétaux et les nutriments.

De plus, notre nomenclature se présente sous la forme substantivale, éliminant par le fait même les verbes (*ailler, brider, fouetter, larder*, etc.) et les adjectifs (*abricoté, aigre, allégé, chocolaté, enfariné*, etc.). Par ailleurs, nous avons conservé les « locutions substantivales » qui décrivent la manière d'apprêter un mets (*à l'anglaise, à la basquaise, à la cantonaise, à la chasseur, à la grecque, à la nage, à l'étuvée, à la marinière*, etc.).

4.1.3 Élaboration de l'arbre du domaine

L'articulation de notre **arbre de domaine**, c'est-à-dire les grandes divisions du domaine, s'est inspirée de deux sources : l'arbre de la *Banque de terminologie du Québec* (*BTQ*) et les grandes divisions de *L'encyclopédie de la cuisine de Jehane Benoit*. La première raison qui motive le choix de ces deux ouvrages est sans aucun doute parce qu'ils représentent bien les différents usages du vocabulaire de l'alimentation, la double appartenance de ce vocabulaire à la fois général et spécialisé. La *BTQ* reflète le côté plus technique, plus spécialisé, alors que *L'encyclopédie de la cuisine*, par sa grande popularité et sa grande diffusion dans les foyers québécois, rend compte davantage des usages de la communauté linguistique. Mais il en existe un second : la complémentarité de présentation de ces ouvrages. En effet, l'arbre de la *BTQ* propose de vastes sous-domaines lesquels sont de nouveau subdivisés en d'autres sous-domaines, alors que *L'encyclopédie de la cuisine* nous offre que de grands génériques. Pour notre étude, la *BTQ* s'avérait trop détaillée, trop terminologique pour nos besoins, et *L'encyclopédie de la cuisine*, trop vaste, trop générale. En confrontant ces deux présentations, nous en sommes venue à construire un arbre à mi-chemin entre les multiples divisions de la *BTQ* et les vastes sous-domaines de *L'encyclopédie de la cuisine de Jehane Benoit*. Ainsi, notre arbre a réuni vingt-six sous-domaines.

4.1.4 Collecte du vocabulaire

La collecte du vocabulaire se fait par dépouillement. La façon la plus pratique consiste à souligner les mots retenus et ensuite à les insérer dans un tableau. Le premier dépouillement constitue la **nomenclature de départ** à laquelle nous greffons d'autres vocables. En privilégiant une démarche circulaire, il faut donc revenir constamment à nos premiers ouvrages dépouillés, afin de vérifier l'oubli de vocables répertoriés lors de dépouillements subséquents, les raisons qui ont motivé leur première exclusion, etc.

4.1.5 Ouvrages dépouillés

Le domaine de l'alimentation constitue un domaine particulièrement intéressant pour la recherche prédictionnairique. Son vocabulaire regroupe à la fois le vocabulaire plus général, connu de tous et le vocabulaire plus spécialisé. Afin d'établir une nomenclature qui soit la plus exhaustive possible, nous avons diversifié la nature des ouvrages dépouillés.

1° Extrait de la nomenclature de deux dictionnaires québécois, le *Dictionnaire du français plus à l'usage des francophones d'Amérique* (*DFP*) 1988 et le *Dictionnaire québécois d'aujourd'hui* (*DQA*) 1993. Le *DFP* compte quelque 60 000 entrées, alors que la nomenclature du *DQA* se situe dans les 40 000 entrées.

Actuellement, selon la délimitation de notre domaine, nous avons recueilli 975 entrées dans le *DQA* et 1313 entrées dans le *DFP*. Naturellement, ce nombre varie légèrement selon le type et la structure du dictionnaire : le *DQA* penche davantage du côté de la polysémie, alors que le *DFP* privilégie plutôt l'homonymie. Pour notre part, nous avons suivi la trace du *DFP*. Par exemple, pour le vocable *blanquette*, le *DQA* attribue deux sens sous une même entrée, soit le « vin blanc mousseux » et le « ragoût de viande blanche », tandis que le *DFP* en fait deux entrées distinctes. Nous avons suivi la démarche du *DFP* en comptant deux entrées pour le vocable *blanquette*. En fusionnant ces deux nomenclatures, nous arrivons à une liste de plus de 1582 entrées

différentes. La fusion de ces deux listes de mots servira de **nomenclature de départ** que nous complétons avec les ouvrages suivants :

2° Extrait d'une nomenclature encyclopédique : le *Dictionnaire encyclopédique des aliments*;

Nous avons choisi un ouvrage à caractère encyclopédique, car nous savions qu'il viendrait enrichir, de par sa nature, les dictionnaires généraux. D'ailleurs, ce dictionnaire encyclopédique québécois décrit plusieurs types d'aliments allant des simples abats aux sucres artificiels en passant par les algues comestibles. Nous y retrouvons autant d'aliments conventionnels que de denrées exotiques. Pour ces raisons, il nous apparaissait pertinent d'ajouter cet outil de référence très complémentaire aux deux précédents. Comme nous l'avons fait pour le *DQA* et le *DFP*, nous avons dépouillé cet ouvrage d'une façon manuelle en utilisant l'index. Pour être encore plus complet, plus précis ou pour une étude plus approfondie des articles, il serait avantageux de « scanner » ce dictionnaire et de l'indexer informatiquement. En dépouillant ce dictionnaire, nous avons retenu un peu plus de 407 entrées, dont une cinquantaine n'apparaissaient pas dans notre nomenclature de départ.

3° Extrait d'une nomenclature terminologique : la *Banque de terminologie du Québec* (*BTQ*);

Nous avons dépouillé la *BTQ* pour plusieurs raisons. D'abord, comme nous l'avons déjà mentionné, le domaine de l'alimentation est à la fois un domaine d'intérêt général et un secteur d'activité spécialisé. De plus, étant un vocabulaire en partie normalisé ou recommandé par les organismes gouvernementaux, l'alimentation doit souvent respecter une terminologie « universelle », notamment dans l'étiquetage des produits, la rédaction des menus, etc. De plus, nous avons privilégié la *BTQ* parce qu'elle donne un aperçu de l'ensemble de nos sous-domaines. Nous avions d'ailleurs choisi cette dernière

comme base de notre arbre du domaine. Il nous semblait alors important de reprendre les quelque 4 000 entrées terminologiques[3] et de les intégrer à notre nomenclature, à titre d'entrées, de sous-entrées ou d'exemples. Notons cependant que parmi ces entrées, plusieurs d'entre-elles se sont avérées trop spécialisées pour être retenues à notre nomenclature (*acétate de calcium*, *achrodextrine*, *acide cholique*, *gomme de xanthane*, *polyvidone*, etc.). Enfin, nous voulions un ouvrage représentant l'organisme qui constitue « l'autorité » officielle en matière de langue publique au Québec, soit l'Office de la langue française. Nous avons retenu 473 entrées « dictionnairiques ».

4° Extrait d'index ou de parties de textes de livres de recettes québécois (*L'encyclopédie de la cuisine de Jehane Benoit*, *Qu'est-ce qu'on mange?* du Cercle des fermières du Québec),

Actuellement, nous avons informatisé une partie du livre de recettes de Jehane Benoit. Naturellement, pour une plus grande précision, pour exploiter davantage la richesse des textes, pour avoir une idée de la fréquence des mots, un dépouillement systématique de l'ensemble de l'ouvrage aurait été préférable. Mais, pour des raisons évidentes de coûts et de temps, cette idée a été réfutée, tout au plus, remise à plus tard.

5° Extrait de mots ou d'exemples dans des textes québécois de la *Banque de données textuelles de Sherbrooke* (*BDTS*).

En utilisant cette banque de données textuelles québécoises, nous completerons la nomenclature établie au cours des quatre premières étapes. Le dépouillement de la *BDTS*, bien qu'elle soit informatisée, nous pose un problème particulier, celui du repérage des

---

3

Mentionnons ici que ces 4000 entrées terminologiques ne constituent pas 4000 entrées de dictionnaire. Par exemple, *boeuf à la mode*, *boeuf au brocoli*, *boeuf aux tomates* représentent 3 formes dans la *BTQ* mais une seule entrée dans un dictionnaire général.

unités lexicales. La *BDTS* compte quelque 5 millions de mots, dont plus de 93 000 d'entre-eux constituent des mots uniques. Il nous apparaît dérisoire d'essayer de repérer le vocabulaire se rattachant au domaine de l'alimentation parmi cette masse énorme de données lexicales, généralement pourvues de plus d'un sens. Nous attriburons donc un rôle différent à la *BDTS*, celui de l'attestation. Ainsi, après avoir complété notre nomenclature, nous chercherons toutes nos entrées dans la banque de données, tant du côté de l'oral que de l'écrit, et nous extrairons les exemples se rattachant à celles-ci. Cette opération nous permettra certainement de compléter certains sens, d'en repérer de nouveaux, de trouver des variantes orthographiques ou phonétiques (*quertons*, *croxignole*, *framboùèse*, etc.), de connaître les entrées les plus fréquentes, en plus d'avoir une idée du vocabulaire de l'alimentation employé au Québec. En extrapollant vers le traitement des données, nous aurons également une banque d'exemples prêts à être utilisés, ainsi que les collocations et les co-occurents de l'entrée étudiée.

## CONCLUSION

Cette recherche avait pour mission première l'établissement d'une nomenclature d'une partie de l'alimentation. Le sujet était d'une envergure colossale : bien des éléments n'ont été qu'effleurés et c'est bien malgré nous. Nous aurions pu étendre notre reflexion jusqu'au traitement des données, mais nous avons dû renoncer, faute d'espace et de temps. Mais comme l'a si bien dit Bernard Quemada (1967 : 3) : « [...] si nous abandonnons dans cet ouvrage nombre de chantiers ouverts, c'est avec l'espoir que leur exploration en sera dès lors facilitée ».

Malgré tout, ce travail nous a permis de réfléchir sur la constitution d'une nomenclature par domaines d'activités et de faire ressortir certains éléments importants.

## BIBLIOGRAPHIE

BENOIT, Jehane. 1991. *L'encyclopédie de la cuisine de Jehane Benoit*. Éditions Mirabel, Saint-Laurent.

BOULANGER, Jean-Claude et autres. 1993. *Dictionnaire québécois d'aujourd'hui*. Montréal, DicoRobert.

DUBOIS, Jean et Claude. 1971. *Introduction à la lexicographie : le dictionnaire*. Librairie Larousse, Paris.

MONETTE, Solange. *Dictionnaire encyclopédique des aliments*. Éditions du Club Québec Loisirs Inc. avec l'autorisation des Éditions Québec/Amérique, Louiseville.

OFFICE DE LA LANGUE FRANÇAISE. *Thésaurus de domaines d'emploi*. Banque de terminologie du Québec, Édition du 21 juin 1990.

OFFICE DE LA LANGUE FRANÇAISE. 1973. *Guide de travail en terminologie*. Gouvernement du Québec, Ministère de l'éducation, 1re version, Cahiers de l'Office de la langue française, Québec.

POIRIER, Claude, Louis MERCIER, Claude VERREAULT. 1988. *Dictionnaire du français plus à l'usage des francophones d'Amérique*. Montréal, CEC.

QUEMADA, Bernard. 1990. « La nouvelle lexicographie », dans *Lingüística aplicada*. Cicle de conferències, Barcelona, pp. 55-78.

QUEMADA, Bernard. 1967. *Les dictionnaires du français moderne, 1539-1863 : étude sur leur histoire, leurs types et leurs méthodes*. Paris.

REY, Alain. 1965. *Les dictionnaires : formes et contenu*. Cahiers de lexicologie, 7, 1965, II, Paris, Éditions Larousse-Didier, pp. 66-102.

ROBERT, Paul. 1993. *Le nouveau petit Robert - dictionnaire alphabétique et analogique de la langue française* (NPR). Paris, Le Robert.

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 305-314.

# L'adaptation d'un dictionnaire d'apprentissage à un public québécois : problèmes et solutions

*Catherine Ouimet*
Université de Montréal

Le présent texte porte sur les dictionnaires d'apprentissage et sur l'adaptation qui en est faite au public québécois. En effet, les dictionnaires d'apprentissage qui sont conçus, pour la plupart, en France doivent être revus et, le cas échéant, modifiés ou complétés avant d'arriver dans les classes québécoises. Nous tenterons donc ici de peindre un tableau général de l'adaptation d'un dictionnaire d'apprentissage dans le contexte québécois. Nous utiliserons, à ces fins, les résultats d'une étude comparative que nous avons effectuée des versions européenne et nord-américaine du *Robert Junior* respectivement parues en 1993 et 1994 (*RJ* et *RJ2*).

## 1. Qu'est-ce qu'un dictionnaire d'apprentissage?

Les dictionnaires d'apprentissage de langue française existent depuis trente ans seulement. Ils ont connu, au cours des dernières décennies, un essor considérable sur le marché. Ces dictionnaires s'adressent aux enfants du primaire et du secondaire. Ils sont de volume réduit : ils ont donc une nomenclature plus restreinte que celle des dictionnaires d'adultes.

Ils sont rédigés dans un langage simple et concret. Les définitions y sont brèves et les mots sont généralement regroupés par famille. Les dictionnaires d'apprentissage sont normatifs, c'est-à-dire qu'ils ne répertorient que les mots considérés «acceptables» par la majorité; les mots fautifs ou vulgaires sont donc évités. Ces ouvrages sont aussi synchroniques, ils ne contiennent ni archaïsmes ni sens anciens. Dictionnaires de langue, ils ne contiennent pas non plus de sens trop techniques ou littéraires.

En revanche, les dictionnaires d'apprentissage comportent très souvent des illustrations — de couleurs ou non — ainsi que des planches thématiques (qui portent sur les papillons, les bateaux, les champignons, etc.). Ces dictionnaires sont conçus pour faciliter la consultation dictionnairique de l'enfant en rendant les informations qui y sont consignées plus accessibles. Plusieurs (comme Josette Rey-Debove (1989 : 19), par exemple) croient cependant que les moyens utilisés pour le faire ne sont pas tout à fait efficaces. Par

exemple, la brièveté de la définition peut contribuer à embrouiller le message qui est destiné à l'enfant.

### 1.1 En France

En France, la popularité de ce type de dictionnaire s'est accrue depuis les années 1960. On y trouve des titres tels que le *Micro-Robert*, le *Dictionnaire du français vivant* chez Bordas et le *Robert Dictionnaire d'aujourd'hui*, qui s'adressent aux étudiants du niveau secondaire. Pour le primaire, on compte, entre autres, le *Maxi-débutants* de la maison Larousse, le *Mes 10 000 mots* paru chez Didier et le *Robert des jeunes*.

### 1.2 Au Québec

Au Québec, la situation lexicographique n'étant pas la même, et la création dictionnairique y étant rare, les dictionnaires d'apprentissage utilisés sont ceux qui sont élaborés en France. Toutefois, le ministère de l'Éducation du Québec exige que ces dictionnaires soient d'abord adaptés aux enfants québécois. C'est le cas notamment du *Larousse maxi-débutants*, du *Dictionnaire CEC Jeunesse* et du *Robert Junior*, que nous traiterons un peu plus loin.

## 2. L'adaptation d'un dictionnaire d'apprentissage

En effet, le ministère de l'Éducation du Québec demande que les dictionnaires faits à l'étranger et qui sont destinés aux classes québécoises soient adaptés à cette clientèle. L'adaptation signifie qu'un dictionnaire a subi des modifications, des suppressions, qu'il a été revu et complété. Les changements opérés peuvent être de nature linguistique, culturelle et méthodologique. L'adaptation a pour but de donner au dictionnaire un contexte qui est connu de ses destinataires éventuels afin qu'ils s'y reconnaissent. La préface de la première version du *CEC Jeunesse*, parue en 1982, donne la justification suivante quant à l'adaptation de ce dernier (Darbelnet, 1982 : 4) :

> Les allusions à la vie française qui ne sont pas utiles à l'usage québécois, ou qui risqueraient de prêter à confusion, ont donc été québécisées. Il eut été absurde, pour ne prendre que l'exemple le plus simple, de garder des prix libellés en francs, alors que le dollar est la monnaie du Canada.

Cependant, l'adaptation ne saurait être trop draconienne, laissant aux usagers de la version modifiée l'occasion de découvrir des réalités différentes des leurs.

## 2.1 L'adaptation au public québécois

En ce qui concerne le Québec, l'éveil à la lexicographie purement québécoise ne se fait qu'à partir des années 1960. Les Québécois ne possèdent pas de dictionnaire propre. Toutefois, ils demandent que les particularités de leur vocabulaire soient répertoriées dans les ouvrages français déjà existants. Une fois ce point gagné, et l'affirmation politique des Québécois se faisant de plus en plus évidente, ceux-ci considèrent que les dictionnaires d'apprentissage devraient, pour leur part, refléter aussi la société québécoise, afin que l'enfant s'y identifie.

La plupart des dictionnaires d'apprentissage sont donc adaptés aux enfants québécois — activité qui prend un essor au cours des années 80, parallèlement à l'élaboration croissante de ce type de dictionnaires.

La Loi québécoise a aussi contribué à cet essor. En effet, l'article 447, paragraphe 3°, de la Loi sur l'instruction publique, donne au gouvernement le pouvoir d'établir un régime pédagogique et ce régime peut, en outre, «déterminer des règles relativement aux manuels scolaires, au matériel didactique ou aux catégories de matériel didactique et à leur accessibilité». Toutefois, ce pouvoir est restreint par l'article 449 qui dit que, en vertu de l'article 22 de la Loi sur le Conseil supérieur de l'éducation, des comités catholiques et protestants ont le pouvoir de décision sur tout ce qui concerne l'enseignement moral et religieux, catholique ou protestant. L'article 23 de la même loi permet aussi à ces comités de donner leur avis au ministre de l'Éducation sur les programmes, les manuels scolaires et le matériel didactique pour l'enseignement autre que l'enseignement moral et religieux. Enfin, l'article 462 de la Loi sur l'instruction publique, mentionnée précédemment, complète ces dispositions en prévoyant que :

> le ministre peut établir la liste des manuels scolaires et du matériel didactique ou des catégories de matériel didactique approuvés par lui et, le cas échéant, par le comité catholique et le comité protestant [...].

### 2.1.1 Modalités d'approbation du Ministère

L'évaluation, qui détermine si un ouvrage adapté est approuvé ou non, est effectuée par le Bureau d'approbation du matériel didactique. Il faut dire que cette évaluation est précédée d'un travail de concertation entre la maison d'édition et le Bureau d'information et de consultation durant la rédaction même de l'ouvrage adapté.

Toutefois, il est important de mentionner le fait que le ministère de l'Éducation ne possède pas, à l'heure actuelle, d'ouvrage précis recueillant les différents critères d'évaluation. Les évaluateurs se basent plutôt sur les politiques et les ouvrages de références émis par le gouvernement québécois et portant sur différents éléments linguistiques. Ainsi, ils se pencheront notamment sur l'*Énoncé d'une politique relative aux québécismes* de l'Office de la langue française, l'*Énoncé d'une politique relative à l'emprunt de formes linguistiques étrangères* et l'*Avis de recommandation sur la féminisation des titres*. Les évaluateurs sont également appelés à se servir de leur jugement (linguistique et autre) afin de décider du sort d'un mot ou d'un exemple dans la version québécoise d'un dictionnaire étranger. Cependant, un responsable du ministère de l'Éducation nous a informée qu'une grille d'évaluation devra prochainement leur être disponible.

Par ailleurs, il est entendu qu'un dictionnaire d'apprentissage adapté au Québec doit être rédigé dans un français québécois standard. Il doit répertorier des réalités du Québec, du Canada ou de l'Amérique du Nord. Il doit aussi refléter la culture québécoise dans les exemples ou les planches thématiques, etc. Enfin, il ne peut comporter d'exemples ou d'expressions à caractères sexiste, raciste ou violent : résultat du mouvement de rectitude politique de plus en plus présent au Québec et en Amérique du Nord.

### 2.2 L'adaptation est-elle source de problèmes?

Certains lexicologues québécois déplorent le recours à l'adaptation des dictionnaires d'apprentissage, car elle repousse à une date ultérieure l'élaboration d'un dictionnaire purement québécois, fait uniquement au Québec. D'autres dénoncent les erreurs — de nature lexicale, par exemple — trop fréquemment présentes dans ces types de dictionnaires. Pour sa part, Claude Simard déclare que

(1990 : 29) «[p]our les dictionnaires du premier cycle du primaire, on note [...] une certaine déficience due soit à une adaptation québécoise rapide et encore trop européenne, soit à l'absence d'une vision précise et cohérente de la variation linguistique». En effet, des emplois courants au Québec sont oubliés, des réalités françaises ne sont pas marquées comme telles et la distinction entre sens français et québécois n'est souvent pas faite.

Par contre, certains (comme Claude Simard (1990 : 45), par exemple) considèrent que l'adaptation, bien qu'améliorable, constitue un moindre mal — si l'on pense aux dictionnaires d'adultes non adaptés qui réunissent pêle-mêle certains québécismes dont on donne un sens erroné, par exemple. De plus, l'adaptation d'un dictionnaire se révèle fort moins coûteuse en temps et en argent.

Nous avons, aux fins de notre mémoire, fait une étude comparative des deux versions du *Robert Junior*, que nous avons mentionné auparavant.

## 3. Le *Robert Junior*

Tout d'abord, il faut noter que le *Robert Junior* — version adaptée ou non — comporte une nomenclature de 20 000 mots et qu'il s'adresse aux élèves de la troisième à la sixième année du primaire.

### 3.1 Étude comparative du *RJ* et du *RJ2*

La version nord-américaine du *Robert Junior*, qui est parue sous la direction de Bruno De Bessé, en 1994, a été revue, complétée et modifiée. Néanmoins, l'ouvrage destiné aux Québécois reste, quant à son contenu et à sa forme, essentiellement le même que son homologue français.

Pour ce qui est de notre étude comparative, elle a porté sur la lettre *B* en raison du nombre élevé d'entrées qui la composent. Nous croyons, de ce fait, qu'elle est suffisamment représentative du contenu lexicographique de ce dictionnaire.

Nous avons aussi comparé la page couverture, l'avant-propos, les illustrations (vignettes et planches), l'atlas et toutes autres informations connexes et complémentaires fournies, ainsi que le

cahier pédagogique.

### 3.1.1 Lettre *B*

La version d'origine compte 839 entrées[1], sous la lettre *B*, tandis que la version adaptée en contient 853. Le nombre de mots étant différent pour les deux versions, on constate déjà que la version nord-américaine — dont le nombre est plus élevé — a subi des ajouts. Par ailleurs, on a aussi procédé à des suppressions. En effet, les deux versions contiennent chacune 797 (Eur. : 95 %; A. N. : 93,4 %) entrées semblables. Toutefois, la version européenne compte 42 (5 %) mots qui ne figurent pas dans la nomenclature de l'adaptation québécoise — pour sa part, cette dernière comprend 56 (6,6 %) entrées non présentes dans son homologue.

Sur les 56 mots ajoutés à la version nord-américaine, certains sont des réalités spécifiquement québécoises, canadiennes ou nord-américaines telles : BAGEL, BLANCHON et BALLON-BALAI. Sur les 42 mots qui ont été supprimés de la version nord-américaine, on retrouve BOURGMESTRE («maire» en Belgique), qui n'est pas d'une grande utilité pour les élèves québécois. Toutefois, on compte des mots qui ont été rejetés, comme BALDAQUIN, BAISEMAIN et BAN, qui auraient pu leur être utile.

Par ailleurs, sur 797 entrées partagées, on en compte 599 (75,2 %) qui sont exactement semblables. Ainsi, 198 (24,8 %) entrées de la version nord-américaine présentent des différences de tous genres.

En fait, on compte 13 types de différences trouvées dans les 198 entrées de la version nord-américaine. Ces différences touchent de nombreux aspects lexicographiques et linguistiques. En voici quelques-unes : certaines entrées comportent un sens différent, c'est le cas de BACCALAURÉAT et de BRIGADIER, qui représentent des réalités différentes au Québec. Cependant, le mot BARIL, qui est usité de la même façon en France qu'au Québec, s'est vu substituer son sens «petit tonneau» par «unité de mesure du pétrole» en version adaptée. Pourquoi n'avoir pas gardé le même sens? Ou

[1] Il faut noter que nous comprenons, ici, exactement tous les mots de la nomenclature, même ceux qui ne font que renvoyer à un autre mot de même signification (**bibite** renvoie à **bébite**).

pourquoi n'avoir pas mis les deux sens dans les deux versions? Ces questions sont restées sans réponse.

La version adaptée comprend aussi des exemples modifiés. Les modifications sont d'ordre syntaxique, sémantique, lexical et culturel. Ainsi, à l'entrée BAL, on a remplacé «Le 14 juillet, il y a de nombreux bals dans toute la France» par «Les bals du Carnaval de Québec». À l'entrée BAS, c'est «Des bas en nylon» qui cède sa place à «Des bas de nylon». Enfin, sous BORDER, on a substitué «Des platanes bordent la route» par «Des sapins bordent la route».

Toujours en ce qui a trait aux exemples, les prénoms féminins qui y figurent à plusieurs reprises, soit *Flora* et *Julie* en version d'origine, ont été substitués par les prénoms *Ève* et *Sarah* en version adaptée. Cette modification est faite dans l'optique que l'enfant est censé se reconnaître grâce à ces prénoms qui lui sont plus familiers.

Certains prénoms masculins ont aussi été remplacés, mais par des prénoms féminins, cette fois-ci. *Yves* et *Alex* deviennent *Anne* et *Luc* devient *Ève*. Ces prénoms ont été changés dans le but de neutraliser les exemples et d'enrayer ainsi les stéréotypes sexistes qui s'y glissent souvent. On retrouvera donc des exemples tels «Luc et Yves se sont battus» remplacé par «Luc et Anne se sont battus» ou encore «Yves a la bosse des maths» modifié par «Anne a la bosse des maths».

Dans le même ordre d'idées, le pronom masculin «il» cèdera souvent sa place au pronom féminin «elle», en version adaptée. En voici quelques exemples : sous BILLOT, on retrouve «Il a fendu à la hache la bûche posée sur le billot» qui devient «Elle a fendu à la hache la bûche posée sur le billot» en version adaptée. Sous BRAVE, c'est : «Il a été brave pendant la guerre» qui est remplacé par «Elle a été brave pendant la guerre». Enfin, sous BOUTON, on trouve «Elle recoud un bouton», en version d'origine, qui est substitué par «Il recoud un bouton» en version adaptée. Cela découle du mouvement de rectitude politique, que nous avons mentionné plus tôt, qui est de plus en plus populaire en Amérique du Nord et qui vise par le langage à mettre tous les êtres sur le même pied d'égalité.

### 3.1.2 Informations complémentaires

Pour ce qui est des informations connexes et complémentaires, que nous avons aussi étudiées, nous ne traiterons que de la page couverture, qui est assez éloquente par ses changements.

| Version européenne | | Version nord-américaine |
|---|---|---|
| Paon de jour | = | Gélinotte huppée |
| Inuit | = | Joueur de baseball |
| Bateau | = | Avion |
| Zèbre | = | Orignal |
| OPEN sur lecteur laser | = | ø sur lecteur laser |

Dans le cas du lecteur laser, c'est tout simplement le mot «OPEN», qui se trouve sur cet objet en version européenne, que l'on a éliminé en version nord-américaine. Cette modification est, selon nous, fort représentative des deux cultures concernées. En effet, en France, un mot anglais ne semble pas étrange sur la page couverture d'un dictionnaire de langue française. Au Québec, un tel mot est inadmissible.

### Conclusion

Enfin, quant à l'adaptation, plusieurs questions sont restées sans réponse au cours de notre étude comparative. En effet, nous n'avions pas de texte précis du ministère de l'Éducation du Québec sur lequel nous baser pour y répondre. Il faut noter par ailleurs que le but de ce travail n'était pas de juger de la qualité du *RJ* ou de l'adaptation qui en est découlée. En fait, le *RJ* ne servait qu'à exemplifier l'adaptation d'un dictionnaire d'apprentissage en général.

Nous avons tenté, au cours de nos recherches, de trouver des solutions aux problèmes que cause l'adaptation d'un dictionnaire d'apprentissage. Malheureusement, nous avons dû reconnaître que seul un fort petit nombre de celles-ci sont réellement valables.

Par exemple, si adaptation il y a, il faut d'abord que les lexicographes québécois définissent clairement les particularités du français québécois afin d'éviter, entre autres, les erreurs lexicales souvent commises en adaptation. Les lexicographes doivent

également remanier la phraséologie des définitions et des exemples qui restent quelques fois trop près du français hexagonal pour être bien compris des locuteurs québécois. Cette phraséologie peut facilement révéler l'origine du dictionnaire et enraye tout effort de convaincre l'usager que ce dernier est québécois; en effet, ces erreurs rendent la «québécisation» fausse et artificielle.

L'adaptation doit être surveillée de près; elle doit être complète (l'adaptateur est presque censé sentir qu'il crée un nouveau dictionnaire). Il est important que l'adaptation soit effectuée par une équipe de locuteurs natifs (spécialistes langagiers et enseignants), qui peuvent contribuer à l'idiomaticité des textes français québécois, ainsi que de lexicographes chevronnés, afin d'assurer la cohérence du contenu lexicographique en fonction des modifications effectuées (système de renvois, synonymie, etc.).

Néanmoins, nous croyons que la création d'un dictionnaire d'apprentissage purement québécois est la meilleure réponse aux multiples problèmes causés par l'adaptation. Cet ouvrage répertorierait directement la réalité du Québec, à partir du fonds lexical du français québécois standard — s'il s'agit d'un dictionnaire normatif. Ce dictionnaire serait, bien sûr, élaboré par des spécialistes québécois. L'expertise y est, il ne manque que la volonté.

## BIBLIOGRAPHIE

### Dictionnaires

Abenaim, Raymonde & Jean-Claude Boulanger. 1986. *Dictionnaire CEC Jeunesse*. Anjou : Centre éducatif et culturel inc.

Bellefonds, Christine de, Sophie Chantreau & Laurence Laporte. 1993. *Robert Junior*. Paris : Dictionnaires Le Robert.

Bessé, Bruno de. 1994. *Robert Junior illustré*. Montréal : Dicorobert Inc.

Darbelnet, Jean. 1982. *Dictionnaire CEC Jeunesse*. Montréal : Centre éducatif et culturel inc.

Davau, Maurice, Marcel Cohen & Maurice Lallemand. 1972. *Dictionnaire du français vivant*. Paris : Bordas.

Didier, Marcel. 1976. *Mes 10 000 mots; le dictionnaire pour l'école*. Paris : Didier.

Lagane, René. 1986. *Larousse maxi-débutants*. Boucherville : Les éd. françaises.

Lagane, René. 1986. *Maxi débutants*. Paris : Larousse.

Rey, Alain. 1971. *Micro-Robert; dictionnaire du français primordial*. Paris : Dictionnaires Le Robert.

Rey, Alain. 1991. *Le Robert Dictionnaire d'aujourd'hui*. Paris : Dictionnaires Le Robert.

Rey-Debove, Josette. 1991. *Le Robert des jeunes*. Paris : Dictionnaires Le Robert.

**Ouvrages**

Auger, Pierre. 1986. «Énoncé de politique de l'Office de la langue française relatif aux québécismes : présentation». In L. Boisvert *et al. La lexicographie québécoise. Bilan et perspectives*. Québec : Les Presses de l'Université Laval, pp. 67-74.

Loi sur l'instruction publique. 1995. L. R. Q., chapitre I-13.3. Québec : Éditeur officiel du Québec.

Loi sur le Conseil supérieur de l'éducation. 1991. L. R. Q., chapitre C-60. Québec : Éditeur officiel du Québec.

Office de la langue française. 1990. *Répertoire des avis linguistiques et terminologiques*. Québec : Les publications du Québec.

Rey-Debove, Josette. 1989. «Dictionnaire d'apprentissage : Que dire aux enfants?». In A. H. Ibrahim (dir.). *Lexiques*. Paris: Hachette, pp. 18-23.

Simard, Claude. 1990. «Les besoins lexicographiques du milieu de l'enseignement du Québec». In Conseil de la langue française. *Dix études portant sur l'aménagement de la langue au Québec*. Québec : Éditeur officiel du Québec, pp. 25-51.

Julie Auger et Yvan Rose (réd.), (1997). *Explorations du lexique*. Québec: CIRAL, 315-325.

# Particularités du vocabulaire sociopolitique québécois

*Nadine Vincent*
Université de Sherbrooke

Cet article présente la méthodologie utilisée dans le cadre d'un doctorat traitant du vocabulaire sociopolitique au Québec. Ce doctorat s'inscrit dans les travaux du CATIFQ (Centre d'analyse et de traitement informatique du français québécois) portant sur la caractérisation des textes de langue publique québécoise.

## 1. INTÉRÊT DU DOMAINE SOCIOPOLITIQUE AU QUÉBEC

Le Québec occupe une position stratégique en Amérique du Nord aux points de vue historique, politique et linguistique. Pour cette raison, la question de l'identité socioculturelle et linguistique y est continuellement au cœur des débats politiques. L'histoire récente est en effet riche en concepts nouveaux et en redénomination de concepts anciens pour leur donner une nouvelle vie, et ce, pour tout ce qui touche l'identité, le politique et le social. Il suffit d'invoquer les concepts vagues de « société distincte », de « partitionnisme » et de « partenariat » pour constater que la société québécoise est en quête d'une redéfinition de son identité politique, redéfinition qui passe nécessairement par une avalanche terminologique. Cela illustre bien l'indissociable lien qui existe, au Québec, entre les domaines politique et linguistique.

## 2. MÉTHODOLOGIE

Nous nous limitons à l'examen des sens sociopolitiques des mots, en axant notre analyse sur une étude synchronique en contexte. Le vocabulaire sociopolitique est étudié en réseau, c'est-à-dire que chaque mot est comparé et mis en relation avec les mots de même famille ainsi qu'avec ses synonymes et antonymes potentiels.

### 2.1. CHOIX ET PRÉSENTATION DU CORPUS

Nous avons choisi de constituer un corpus à partir des mémoires présentés à la *Commission sur l'avenir politique et constitutionnel du Québec* (Commission Bélanger-Campeau), qui a siégé de novembre 1990 à janvier 1991. Le mandat principal de la Com-

mission était « d'étudier et d'analyser le statut politique et constitutionnel du Québec et de formuler à cet égard des recommandations » (Rapport de la Commission 1991:1).

Compte tenu du contexte dans lequel s'est tenue la Commission Bélanger-Campeau, on peut supposer que les textes qui y ont été présentés contiennent un vocabulaire sociopolitique particulièrement abondant en ce qui touche la question de l'identité du Québec. Il faut aussi noter que le court laps de temps (à peine quelques mois) pendant lequel ont été préparés et présentés ces mémoires permet une unité de synchronie qui assure la cohérence des données. L'intérêt du corpus Bélanger-Campeau est donc qu'il fournit des données pertinentes (sociopolitiques), comparables (synchronie uniforme) et en quantité appréciable.

## 2.2. CONSTITUTION DU CORPUS

Dans l'ensemble des mémoires présentés à la Commission Bélanger-Campeau, nous avons sélectionné 100 extraits de mémoire comptant 2500 mots chacun, pour obtenir un corpus final de 250 000 occurrences.

## 3. ÉTABLISSEMENT DE LA NOMENCLATURE

Il importe de signaler que nous avons choisi de traiter le vocabulaire sociopolitique comme relevant de la langue générale (plutôt que comme un vocabulaire technique ou spécialisé). De plus, le vocabulaire sociopolitique est théoriquement transparent, puisqu'il doit être compréhensible par l'ensemble de la population. C'est donc un domaine pour lequel on ne s'attend pas à trouver un nombre significatif de termes abscons et de notions inaccessibles aux non-initiés.

Nous avons procédé au repérage du vocabulaire sociopolitique à l'aide d'un certain nombre de critères.

## 3.1. FRÉQUENCE

Le premier critère servant à identifier le vocabulaire sociopolitique est bien sûr sa fréquence relative dans le corpus Bélanger-Campeau. La fréquence d'un mot n'indique évidemment pas son

caractère sociopolitique, mais elle permet d'identifier les mots les plus usuels.

### 3.2. COOCCURRENTS ET GROUPES COMPLEXES

Lors de cette analyse du vocabulaire sociopolitique, nous accordons une attention particulière au repérage des cooccurrents. En plus de l'usage syntaxique d'un mot, on identifie ainsi des groupes complexes. En effet, les mots toujours employés ensemble adoptent un nouveau sens que n'a pas chacun des constituants (ex.: *société distincte*, *affirmation nationale*, etc.).

De plus, l'examen des cooccurents permet de cerner pour certains mots, non sociopolitiques à l'origine, un champ sémantique nouveau, parfois propre au français québécois. Par exemple, pour un locuteur québécois, *accession* renvoie notamment à *souveraineté* ou à *indépendance*, alors que *rapatriement* et *enchâssement* sont très souvent associés à *constitution*.

### 3.3. FAMILLES DE MOTS

Il est également utile de regrouper les mots en familles pour faciliter la détection des glissements de sens et des néologismes. Ainsi, certaines séries s'établissent autour des mots suivants :

*fédéral* (*fédéralisme*, *crypto-fédéraliste*, *anti-fédéral*, etc.),
*nation* (*national*, *nationalisme*, *nationalitaire*, etc.),
*souverain* (*souveraineté*, *souverainisme*, *souverainiste*, etc.).

### 3.4. COUPLES ANTONYMIQUES ET SYNONYMIQUES

Dans le même ordre d'idée, les couples de synonymes et d'antonymes sont isolés. Par exemple, *indépendance* et *souveraineté* sont-ils deux concepts interchangeables ou deux notions distinctes? L'opposition à un possible antonyme, comme *fédéralisme*, concourt à consolider les définitions.

### 3.5. ANALYSE MORPHOLOGIQUE

Finalement, la morphologie nous intéresse surtout pour l'analyse des dérivés et des néologismes. Ainsi, des études antérieures portant sur le français du XIX[e] siècle ont déjà identifié quelques

affixes à connotation sociopolitique; cf. Dubois (1962), Rabotin (1975), Jousselin (1977). Il s'agit de préfixes comme : *ultra-*, *archi-*, *extra-*, *lèse-*, *anti-* et *quasi* et de suffixes comme : *-isme*, *-iste*, *-ité*, *-cratie*, *-at*, *-âtre* et *-cide*.

## 4. ÉLABORATION DES DÉFINITIONS

Une fois cernée une première nomenclature, nous passons au traitement du vocabulaire sociopolitique, opération qui s'articule en plusieurs étapes. En guise d'exemple, tentons une première ébauche de la définition du mot *souveraineté*...

### 4.1. RECENSION DES DICTIONNAIRES

La première étape consiste à recenser, dans les principaux dictionnaires existants, autant français que québécois (voir bibliographie), les sens sociopolitiques des mots retenus. Nous groupons les sens et sous-sens en fonction de leurs sèmes dominants.

Pour le mot *souveraineté*, nous isolons deux sens, dont le premier se résume en :

1) une qualité, un principe, parfois abstrait;
2) propre à un gouvernement souverain, à une nation, à une personne physique ou morale dans l'État, dans le corps politique;
3) qui détient le pouvoir, l'autorité suprême.

Quant au deuxième sens, il se présente plutôt comme :

1) un caractère ou une qualité;
2) propre à un État, un organe, un organisme ou à une « étendue de pays »;
3) qui n'est soumis au contrôle, qui n'est sous la dépendance d'aucun autre État, ce qui implique l'exclusivité de sa compétence sur le territoire national et son indépendance dans l'ordre international.

### 4.2. DÉPOUILLEMENT DU CORPUS

Le dépouillement du corpus Bélanger-Campeau constitue l'étape fondamentale de notre analyse, puisqu'il sert de base à

l'ébauche des définitions. Ces définitions sont construites en fonction de l'usage des mots en contexte — ce qui implique leurs relations de cooccurrence — mais aussi en parallèle avec les mots de même famille et en confrontation avec les synonymes et antonymes potentiels.

### 4.2.1. LE MOT *SOUVERAINETÉ*

L'analyse des 278 occurrences du mot *souveraineté* contenues dans le corpus Bélanger-Campeau confirme l'existence des deux sens présentés dans les dictionnaires courants (« principe » et « application concrète du principe »). Cependant, les définitions existantes manquent de précision et s'appliquent parfois difficilement au contexte québécois.

Le besoin de définitions construites en fonction du contexte québécois s'illustre notamment par l'analyse des cooccurrents, c'est-à-dire des termes habituellement utilisés avec le mot *souveraineté*. En effet, en regroupant les diverses utilisations du mot *souveraineté*, un découpage se fait naturellement. Cela nous permet d'en préciser l'usage syntaxique. On peut, par exemple, *accepter, assumer, choisir, détenir, faire, revendiquer, souhaiter, vouloir la souveraineté*.

Mais, fait plus intéressant d'un point de vue sémantique et dictionnairique, l'analyse de la cooccurrence permet de mieux cerner les différents sens sociopolitiques du terme étudié.

**a) Sens 1**

On rencontre le *concept*, l'*idée*, la *notion abstraite* de *souveraineté*, la *question*, le *sujet*, la *thèse* de la *souveraineté*, usages qui correspondent bien sûr au premier sens mentionné dans les dictionnaires. Mais plus encore, on identifie, grâce à l'analyse des cooccurrents, les détenteurs potentiels de la souveraineté. Ainsi, le complément du nom *souveraineté*, sens 1, précise QUI détient la souveraineté. Le corpus nous propose les exemples suivants :

> *souveraineté du peuple, populaire; souveraineté nationale; souveraineté nationale canadienne; souveraineté nationale du Québec; souveraineté du Québec, québécoise; souveraineté du gouvernement du Québec.*

1a) *La reconnaissance de la souveraineté populaire est pour nous un des premiers pas à franchir dans l'accession à l'indépendance.*
1b) *Le gouvernement fédéral [...] est le seul porte-parole autorisé de la souveraineté nationale canadienne.*

Dans le corpus Bélanger-Campeau, le terme *souveraineté* utilisé en emploi absolu désigne la souveraineté du Québec.

1c) *Selon certains, la souveraineté va briser le Canada actuel.*

**b) Sens 2**

Le sens 2 présente des difficultés particulières. D'abord, dans la presque totalité des définitions existantes (sauf celles du TLF et du PLI), la souveraineté est définie par rapport à un autre État, ce qui ne trouve aucun écho dans le corpus. En effet, la souveraineté y semble plutôt émerger d'une entité détentrice du pouvoir (peuple, gouvernement). De plus, l'examen des cooccurrents révèle une nuance importante dans ce deuxième sens du mot *souveraineté*. Accompagné d'un cooccurrent, ce mot signifierait : « pouvoir exclusif d'une autorité dans un champ de compétence donné », ce champ de compétence étant identifié par le cooccurrent en question. On parle ainsi de *souveraineté culturelle, économique, politique, territoriale*, c'est-à-dire d'un pouvoir exclusif sur la culture, l'économie, la politique ou le territoire. Citons des exemples tirés du corpus:

2a) *Bien entendu, la question de la souveraineté culturelle réclamée par le Premier ministre Bourassa dès 1972 doit être débattue.*
2b) *Le Québec devrait rapatrier tous les pouvoirs et mettre en marche démocratiquement et rapidement les processus conduisant à la souveraineté politique, voire économique.*
2c) *Or, il est impérieux de reconnaître qu'il ne saurait y avoir de souveraineté politique sans souveraineté du territoire.*

Par ailleurs, le mot *souveraineté* employé seul fait référence à une exclusivité des pouvoirs sur l'ensemble des champs de compétence.

2d) *Et, la souveraineté, plusieurs l'expriment clairement, c'est d'abord et avant tout le rapatriement complet et unilatéral de tous les pouvoirs de décision, la prise en charge complète de notre autonomie politique.*

Cette nuance sémantique capitale ne peut être exprimée sans le recours aux cooccurrents. De fait, ce dernier sous-sens est confirmé abondamment à l'aide d'exemples supposant un possible fractionnement de la souveraineté : *souveraineté pleine et entière, totale, complète*, etc.

En fonction du corpus Bélanger-Campeau, le terme *souveraineté* correspond donc plutôt aux notions suivantes :

**sens 1** Principe de pouvoir, d'autorité suprême.
**sens 2** Pouvoir exclusif d'une autorité (peuple, gouvernement, État) sur un champ de compétence donné.
(en emploi absolu) ♦ Pouvoir exclusif d'une autorité (peuple, gouvernement, État) sur tous les champs de compétence touchant son territoire et sur ses relations internationales.

### 4.2.2. SÉRIE ÉTYMOLOGIQUE

Pour consolider la définition, il faut encore vérifier la cohérence avec les mots de même racine étymologique. Ainsi, si le sème de « l'exclusivité du pouvoir » est indissociable du mot *souveraineté*, il devrait en être de même pour les mots *souverain, souverainiste, souverainisme* et *souverainement*. Nous abordons ici le cas du mot *souverain*.

Cet adjectif est ainsi défini par tous : *qui exerce un pouvoir suprême, qui, dans son domaine, n'est subordonné à personne, qui possède la souveraineté*. Cependant, le TLF nuance en incluant des notions de soumission potentielle selon qu'il s'agisse d'un État, d'une personne morale ou d'une collectivité politique. Ainsi, les deux premiers détiennent la souveraineté, le pouvoir suprême, alors que la troisième, la collectivité politique, *se gouverne soi-même tout en pouvant relever d'une autorité supérieure* (TLF). Cette nuance ne correspond pas aux exemples du corpus Bélanger-Campeau ni aux

définitions retenues du mot *souveraineté*. De plus, le peuple est identifié par le TLF à une personne morale plutôt qu'à une collectivité politique.

Il nous semble donc justifié de définir ainsi le mot *souverain*: « Qui possède et exerce la souveraineté (sens 2), en parlant d'un État, d'une collectivité politique ou d'une personne morale. »

### 4.2.3. SYNONYMES ET ANTONYMES

Il est essentiel de comparer la définition du mot étudié à celles de ses synonymes et antonymes potentiels. Procédons à une comparaison de *souveraineté* avec son plus proche concurrent synonymique : *indépendance*.

**a) Comparaison des définitions existantes**

Les dictionnaires Robert (GR, NPR, DQA) définissent les deux termes de façon similaire, en faisant appel à une citation de Henri Capitant (1865-1937), juriste français auteur d'ouvrages de droit civil.

| SOUVERAINETÉ | INDÉPENDANCE |
|---|---|
| Caractère d'un État, d'un organe « qui n'est soumis à aucun autre État ou organe » (Capitant) | « Situation d'un organe ou d'une collectivité qui n'est pas soumis à un autre organe ou à une autre collectivité » (Capitant) |

La différence entre ces deux termes se résume donc au fait que la souveraineté est le « caractère d'un État », alors que l'indépendance serait plutôt la « situation d'une collectivité »...

Notons que le même parallèle, bien que moins littéral, s'établit également pour les dictionnaires Larousse et pour le Bélisle. Ces dictionnaires présentent donc ces vocables comme étant de parfaits synonymes. Quant au DFP, il centre ses définitions sur les adjectifs *souverain* et *indépendant*.

Seul le TLF présente une analyse différente en établissant une distinction tangible entre les termes *souveraineté* et *indépendance*.

| SOUVERAINETÉ | INDÉPENDANCE |
|---|---|
| Qualité propre à l'État qui possède le pouvoir suprême (impliquant la souveraineté interne, externe et l'indépendance vis-à-vis des puissances étrangères). | « Situation d'un organe ou d'une collectivité qui n'est pas soumis à l'autorité d'un autre organe ou d'une autre collectivité » (CAP. 1936) |

Deux remarques s'imposent ici. D'abord, le TLF définit la souveraineté comme étant la propriété d'un État, alors que l'indépendance s'établit en fonction d'un autre État. Ensuite, la distinction entre les deux termes est assez nette pour que le mot *indépendance* serve à définir le concept de « souveraineté ».

**b) <u>Comparaison des contextes du corpus</u>**

Une première observation des cooccurrents dans le corpus Bélanger-Campeau indique que l'usage des mots *souveraineté* et *indépendance* semble en tous points similaire. Ainsi, on peut *accéder à, choisir, déclarer, être en faveur de, faire, prôner* aussi bien la *souveraineté* que l'*indépendance* et ces mots désignent tous deux un concept, une idée ou une question.

La nuance principale en est une de connotation et apparaît dans la gradation que l'on propose des possibles statuts constitutionnels pour le Québec. On y remarque alors que l'indépendance désigne le pôle extrême de division entre le Québec et le Canada.

3a) *le fédéralisme est en soi un compromis entre l'union complète et l'indépendance.*

3b) *[...] lequel des régimes politiques suivants favorisez-vous? Le maintien du régime fédéral actuel... L'accession du Québec à l'indépendance... La souveraineté du Québec avec une association au Canada...*

Plus souvent que le mot *souveraineté*, le mot *indépendance* est employé pour nommer une séparation complète, une coupure entre le Québec et le Canada. On ne suppose pas une autarcie, mais bien une absence de lien avec le Canada, de quelque ordre que ce soit. Il serait alors juste de supposer qu'employé en absolu, le terme *souveraineté* désigne « la possession de l'exclusivité des pouvoirs dans tous les champs de compétence », alors que le terme *indépendance* désigne plutôt « l'absence de soumission à une entité extérieure ». Nous rejoignons donc ici l'analyse du TLF.

Par ailleurs, si cette non-soumission, incluse dans le concept d'« indépendance », s'applique, en absolu, à tous les champs de compétence, elle peut être limitée, tout comme l'application de la souveraineté, à des champs particuliers.

3c) *[...] faire l'indépendance politique sans la séparation économique* (l/m(360)Charbonne 1)

3d) *Le parlement déclare l'indépendance culturelle des Québécois et exige sa pleine souveraineté dans tous les domaines qui assurent son avenir comme terre française en Amérique.* (l/m(036)JeSimoneau 10)

Notons enfin que tout comme l'examen des cooccurrents dans le corpus Bélanger-Campeau permet d'identifier les détenteurs de la souveraineté, il permet également d'indiquer les bénéficiaires de l'indépendance (*du Québec, nationale, des Québécois*).

**CONCLUSION**

Voilà qui complète un premier aperçu de la méthode envisagée pour traiter le vocabulaire sociopolitique au Québec. Il est encore trop tôt pour établir une définition précise du mot *souveraineté*. Il faudrait encore le comparer à d'autres synonymes, puis à des possibles antonymes.

Nous espérons tout de même avoir démontré l'intérêt de la définition d'un vocabulaire en réseau, de l'utilisation d'un corpus synchronique et de l'étude du vocabulaire sociopolitique québécois, de même que la richesse de l'étude des cooccurrents et le besoin de définitions élaborées en fonction d'un contexte québécois.

Cette méthode d'analyse sera bien sûr nuancée et raffinée au cours des recherches à venir.

**BIBLIOGRAPHIE**

Commission Bélanger-Campeau. 1991. *L'Avenir politique et constitutionnel du Québec*, rapport de la Commission, Québec: Publications du Québec, 180 p.

Dubois, Jean. 1962. *Le vocabulaire politique et social en France de 1869 à 1872*, Paris: Librairie Larousse.

Jousselin, Jean-Pierre. 1977. *Vocabulaire politique en usage en 1867 dans «L'Union des Cantons de l'Est» et «Le Défricheur»*, Mémoire de maîtrise, Université de Sherbrooke.

Prost, A. 1969. «Vocabulaire et typologie des familles politiques», *Cahiers de lexicologie*, XIV(1):115-126.

Rabotin, Maurice. 1975. *Le vocabulaire politique et socio-ethnique à Montréal de 1839 à 1842*, Montréal:Didier.

Dictionnaires recensés

BEL Bélisle, Louis-Alexandre. 1971. *Dictionnaire général de la langue française au Canada*, Montréal: Bélisle Éditeur Inc. (deuxième édition)

DFP *Dictionnaire du français plus*, à l'usage des francophones d'Amérique, 1988, Montréal: CEC.

DQA *Dictionnaire québécois d'aujourd'hui*, 1993. Montréal: Dicorobert.

GLE *Grand Larousse encyclopédique*

GLLF *Grand Larousse de la langue française*, 1971-1978, Paris: Larousse, 6 vol.

GR *Dictionnaire alphabétique et analogique de la langue française*, 1985, Paris: Le Robert, 9 vol.

NPR *Nouveau petit Robert, dictionnaire alphabétique et analogique de la langue française*, 1993, Paris: Le Robert.

PLI *Petit Larousse illustré 1995*, Paris: Larousse.

TLF *Trésor de la langue française : dictionnaire de la langue du XIX^e^ et du XX^e^ siècles (1789-1960)*, 1971-1974, Paris: Éditions du CNRS.

*Pour la première fois de son histoire, l'**Association canadienne-française pour l'avancement des sciences (ACFAS)** a tenu son congrès annuel à l'Université McGill. C'est dans ce cadre que s'est déroulé un colloque intitulé* **Explorations du lexique**, *organisé par Julie Auger et des collaborateurs des universités McGill et de Sherbrooke.*

*Ce livre regroupe les meilleurs textes qui ont été présentés au cours de ce colloque et qui ont été acceptés par le comité de rédaction. L'ouvrage est divisé selon cinq sections thématiques: la représentation mentale du lexique bilingue, la sémantique lexicale et la pragmatique, les aspects sociolinguistiques du lexique, les emprunts et les lexiques spécialisés.*